

GenCore version 4.5
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 16860.4 seconds
 (without alignments)
 12121.202 Million cell updates/sec

Title: US-09-700-906A-L_COPY_197_9962
 Perfect score: 9766
 Sequence: 1 atdttgcccacgagacgcct.....tcataggacagtgaagata 9766

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_on:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	9766	100.0	12493	6	AX009576
2	9715	99.5	12515	6	AX336950
3	9715	99.5	12515	9	HSMKI67
4	8286	84.8	11435	9	HSMKI67A
5	6847	70.1	12965	9	HSAMAK67
c 6	5227	53.5	123057	9	AL390236
c 7	2855	29.2	140466	9	AL355529
8	629	6.4	733	6	AX147668
9	456	4.7	568	6	AX341136
c 10	245	2.5	318	11	G33252
c 11	199	2.0	332	11	G03247
c 12	198	2.0	409	6	AX336882
c 13	110	1.1	158	6	AX072812
14	93	1.0	14041	9	HSKI67
c 15	64	0.7	176433	9	HS71L16
16	55	0.6	391	9	MMKI67391
17	51	0.5	51	6	AX204486
18	40	0.4	430	9	MMKI67430
19	28	0.3	9471	10	MMKI67
20	24	0.2	35540	8	SPBC36
c 21	24	0.2	63170	10	AL645799
c 22	24	0.2	161997	2	AL645526
23	24	0.2	194575	2	AC023140
24	24	0.2	229232	2	AC096624
25	24	0.2	247899	2	AC025909
c 26	23	0.2	23	6	AX009578
27	23	0.2	184	4	BOVMICREB
28	23	0.2	53419	10	MMU35323
c 29	23	0.2	62280	2	AC094189
30	23	0.2	71299	2	AC107772
c 31	23	0.2	73198	9	HSJ127L3
32	23	0.2	73402	2	AC101535
33	23	0.2	73886	2	AC106156
c 34	23	0.2	78693	9	HSDJ475B7
35	23	0.2	96951	9	AC026433
36	23	0.2	100653	2	AC093215
c 37	23	0.2	124814	9	AL161436
c 38	23	0.2	125630	2	AC097551
39	23	0.2	145157	2	AC103108
40	23	0.2	148000	2	AC108922
41	23	0.2	155397	9	AL590523
42	23	0.2	156662	2	AC024454
43	23	0.2	157639	2	AC103026
44	23	0.2	164562	2	AC027471
45	23	0.2	165700	2	AC099450

ALIGNMENTS

RESULT	1				
AX009576					
LOCUS	AX009576	Sequence 1	12493 bp	DNA	linear
DEFINITION	Sequence 1	from Patent	WO9961607.		
ACCESSION	AX009576				
VERSION	AX009576.1	GI:9996814			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				

ORGANISM	homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 12493)
AUTHORS	Deinert, I., Boehle, A., Gerdes, J. and Flad, H.D.
TITLE	Antisense oligonucleotides for treating proliferating cells
JOURNAL	Patent: WO 9961607-A 1 02-DEC-1999; DEINERT IRINA (DE); BOEHLE ANDREAS (DE); GERDES JOHANNES (DE); FLAD HANS DIETER (DE); FORSCHUNGSZENTRUM BORSTEL ZENT (DE)

FEATURES

Source

1. 12493

/organism="Homo sapiens"

/db_xref="taxon:9606"

197. .9967

/note="unnamed protein product"

```

/codon_start=-1
/protein_id="CAC07498.1"
/db_xref="gi:9996815"
/translation="MWPTRLRVLTIKRSVGDPHPPLSLSTLCFGRGTECDRIQLPQV
SKOCHKVEHQEALTNHFSNTNPVNGSVIDEPVRLKHGCDVTLIDRSFVENE
QNGKSTFPPKRIQDEPARRVSRSSSDDEKAQSKAYSKITEGKVGSGNQVQVHI
NKVEDTADDSKDSVAQGTIIVHSEHAGRNAAPIQSGDFKEISLVKLSRYGEL
KSVPTQCLDINSKNSFPWKLIVSEVKLELDVKQKENVLQYCRKSGLDIATYAK
ADLOGETQLLVSRKSRPGSGSHVAAEPAQELDQNGKGRDVSQVTSKAVG
ASFPLYEPAKWKTVOYSGOQNSQKHKNKDLVYTGRRSENVLGKSGEGFAGDKTLP
RKLSTRNTPAKVEDAADSATKPENLSKTRGSIPTDVEVLPTETEIHNEPFLTLMT
QVERIKQDSLQKELGTTCQWCMGSLPGLUSSVDIINFGDSINESGSIPLKRRRV
GHEHLPELDENLPKPTLKGAEPTKRKSLVMTTPVLKKIIEKQSGQSGSGSE
THVEYKASQVISPAPSRKTPVAGSDORRRSCKTAPASSKQTEYPTGGRGERTV
LOKRVYSIKRSQHDILQTKSKRRSGSEANLIVAKSWADVYVKGAKGTQTKVILKHGPQ
RSMNKRORPATPKKPVGEVHQSFTGHANSPTLIIGKAETEKVHPVAPRYVLNPP
ISNQMKDFKDLGSIANFKPTVKEQQLQSTCHIASNSENLLGKQFGQDSGEPL
LPTSSFGNVFFSAQNAKQSDKCSASPLRRQCIRENGNVAIDPRNTYKMTSL
KTSDETPSTKSTVTVNNGSRSETRFNKQIPVSKSEETNEIVEICILRGQKATLL
QORRGEMKEITRPFETYKENIELKENDKMKAMKRSRTWGQCAPMSDLTDLKSLPD
TELMADTARGONLLQTODHAKAPKSEKGIKMPQCSLOPEPINPTHTQQLKASLG
KVGVEELLVAVGTRTQTHREBPAGDGKSIPTFKESPQIOLDPAARYVGMKKW
PRTPEEAQSLDLGAFKELFQTPGPGSESMDEKTTIAKSSPPRESVDTPSTKOW
PKRSKADVDEEELALRKLTPSAGKAWLTPKAGDEKIDKAFMGTPVQKLODGLFT
PGSKRQLOTPKQAQALDLGAFKELFQTPGHEELVAAGKTTIKPCDSQSDPQVPT
TSTQKPKRSIRKADVEGELLACNLNMP SAGKAMHTPKPSGEEKDITIIFVGTQV
DLTENLTKSRRPOTPEBEAQALEDLGFELFQTPGHTTEAVAAGTKTIIPGCCSP
ESADTPSTRQPKTPLEKRDQKLSALKKITQTSGETTHTDKVPGEDEKSIINAPRE
TAKQKLDPAASVYSGSKRHPKTEKAQPLEDLAGMKELFQTPVCTDKPTTHEKTKI
RSQDPDTPSTSSQPSKRSKRVDEEEFALRRKPTSGAGKAMHTPKPAVSSEKNTI
AFMGTVPQKLODLNLTKSKRRRLTSLGKQVLEELAVGLKELFQTSGETTHTE
AKVAKSSQDPLDKNLPASSKRRRLTSLGKQVLEELAVGLKELFQTSGETTHTEPT
DKSKMFKPVSQIOLDASAALSTGSKRQLRTPCKGSEVPEDLAGFIELFQTPSH
MTNEKTKVSVRASQDPLDTPSTSKQPKRSLBKADTEBEFLAFKQTPSAGKAMHT
PKPAVGEKIDINTPLGTPVQKLDQPGNLPGSNRRLOTRKEKAQLEBTLFQTP
CTDNPTADEKTTKILCKSPQSDPADTPNTKQRPKRSLLKADVEEELFAPRLTPSA
GKAMHTPKAAVGEKIDINTFVGTPEVKLDLLGNLPGSKRRPQTPKEKAKALEDLAG
ELFQTPGHTSEMDTKITVEYCSKSPQDPQKMDPNYQTKRQKLSISLWGVKEEVLPG
KLTOTSGTKTOTHRETAGDKSIKAFKESAKQMDPNYGTGMERWPTPKERAQSL
DLAGKELFQTPDHTSESTDKTKIACKSPPEMDTPTSTRRPRLPLGKRDIVE
ELSAKLQITQTHTDKVPGEDEKGINVRETAKQDLPAAVYSGSKQKQPTPKGAQ
LEDLAGLKEFQTPDKPTTHKTIACRSQDPQVGTPTIIPQSKRSRLKQAP
EESLALBKRRTPSVGKAMDTPKAGDEKMKAFMGPVQKLODLGNLPGSKRWQPT
KEKAQLEDLAGFELFQTPGOTQTHDEKTKIACKSPQDPDTPASTQKRPKRL
KQADVEEFLALRKRTPSAGKAMDTPKPAVSDEKNIITFVETPVQKLDLLGNLPGSK
RQDTPKEKAQLEDLGAFKELFQTPGHTSEMDTKITVEYCSKSPQSEFKTSRSK
RLKILPVKVMKEEPLAVSKLTRISGETTQTHTEPTGDSKSIKAFKESQIOLDPAAS
VTGSRQRQTRKEKARALEDLDFKELFQTSQTHTEHMTIDKNTKIPCKSPPELDT
ATSKRQKTPTRKVEKEELSAVERLTQSGQTHTEHPASGDEGIKLPGKRAKKPK
NPVEEPPRRRPAPKEKAQPLEDLAGTTELSTSGHTQSLTAGKATKIPCESPPL
VDDTTAKTRHLTRVOKQVYKEPSAVKTTQTSGETTDADKEPAGDKGKALKESA
KOTPAVASVYSGRRRPAPRESQAQLEDLAGKQDPAAGHTSEMDTKTKIIPCKSS
PELEDTASVYSGRRRPAPRESQAQLEDLAGKQDPAAGHTSEMDTKTKIIPCKSS
QPAKNVDAEDVIGSRQPRAPKEKAQPLEDLAFSQBLSQTPGHTTELANGADSTFS
APKQTPDGSGLPKLTSRRVRLPVPKGVVYTRDPVKYSQSKNSLPLLPFRKGKG
DCSGVTGTRKRLCMAPEIIVBELPAGKQVRVAPRARGKSEPPVIMKRSRLTKSRIE
PAEELNSMDKTNKEHKLODSVPEKISILSRSRQDKTEAQOITBEVFLASRIET
RNEKPKMTPSPMDIQNDGARKPIPRDKVTENKRCILSRQNGESSQPKVAESGEG
KSAYLQMQKQKGEAGNSDMSCLRSRKTQSQPAASTLESKSVQRVTRSVKRAENPK
KAEDNVCKKTIITTHSRGSEDI"
BASE COUNT      4143 a 3048 c 2329 g 2373 t
ORIGIN

Query Match      100.0%; Score 9766; DB 6; Length 12493;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9766; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgccccagagcgctgttactatcaaaagagcgaggtcgagcggtcccacttt 60
|||||
Db 197 ATGTGCCCCACAGAGCGCTGTTACTATCAAAAGAGCGGGGTGCACGGTCCCACTTT 256
|||||

QY 61 cccctgagcctcagacacctgtgttttgggaagggtattgaatgtgacatccgtatccag 120
|||||

```

Db 1337 GAAGGCTTCAGGCTGTGTGATAAAACATCTTACTCCAGGAAGCTTCAACTAGAAATCGA 1396
Qy 1201 acaccagctaaagtgaagatgcagctgactctgccactaagccagaaaaatctctcttcc 1260
Db 1397 ACACCAGCTAAAGTTGAAGATGCAGCTGACTTGCCTACTAAGCCAGAAAATCTCTCTCC 1456
Qy 1261 aaaccagagaagtattctctacagatgtggaagttctgctactcgaaactgaaattcac 1320
Db 1457 AAAACCCAGAGGAAGTATTCCTACAGATGTGGAAGTTCTGCCTACGGAAACTGAAATTCAC 1516
Qy 1321 aatgagccatttttaactctgtgctcaactcaagttagaggaagatccaaaaagattcc 1380
Db 1517 AATGAGCCATTTTAACTCTGTGCTCCTCACTCAAGTTGAGAGGAAGATCCAAAAGGATTC 1576
Qy 1381 ctacgcaagcctgagaattgggcaactcacagctgacagatgtgctctgggttacctgggt 1440
Db 1577 CTCAGCAAGCCTGAGAAATTTGGGCACATACAGCTGACAGATGTCTCTGGGTTACCTTGGT 1636
Qy 1441 cttagttcagttgatatacaaaactttggtgattccattaatgagagtgagggaatacct 1500
Db 1637 CTTAGTTTCACTGATATCAACAACCTTTGGTGATTTCCTAATGAGAGTGAGGGAATACCT 1696
Qy 1501 ttgaaaaagaaggcgtgtgctcttgggtggcgcacctaagacctgaactatttggatgaaac 1560
Db 1697 TTGAAAAGAAGGCGTGTCTCTTTGGTGGCCACCCTAAGACCTGAACATATTTTGATGAAAC 1756
Qy 1561 ttgctccttaatacgcctctcaaaaggggagaagccccaacaaagaaagttctctggtta 1620
Db 1757 TTGCTCTCTAATAGCCCTCTCAAAAGGGGAGAGCCCCCAACCAAAAGAAAGTCTCTGGTA 1816
Qy 1621 atgcacactccactgtctctgaagaaatcatcaaggaaacgcctcaacatcaggaaaa 1680
Db 1817 ATGCACACTCCACCTGTCTCTGAAGAAATCATCAAGGAACAGCCCTCAACCATCAGGAAA 1876
Qy 1681 caagatcaggttcagaaatccatgtggaagtgaagcacaagaacttgggttaataagccct 1740
Db 1877 CAAGATCAGGTTCCAGAAATCCATGTGGAGTGAAGGCACAAAGCTTGGTTTAAAGCCCT 1936
Qy 1741 ccagctcctagtctagaaaaactccagttgcccagtgatcaacgccgtagctctgcaaa 1800
Db 1937 CCAGCTCTAGTCTAGTCAAAACTCCAGTTGCCAGTGATCAACGCCGTAGGTCCTGCANA 1996
Qy 1801 acagccctccttcagcagcaaatctcagacagaggttctctaagagagagggagaaga 1860
Db 1997 ACAGCCCTCTGCTCCAGCAGCAAAATCTCAGACAGAGGTTCTTAAGAGAGGAGGAGAA 2056
Qy 1861 gtggcaacctgcttcaaaagagagtgctctatcagccgaagtcacaatgatattttacag 1920
Db 2057 GTGGCAACCTGCTTCAAAAGAGAGTGTCTATCAGCCGAAGTCAACATGATATTTTACAG 2116
Qy 1921 atgatatgttccaaaagaagaagtgggtgcttcggaagcaaatctgattgttgcaaaatca 1980
Db 2117 ATGATATGTTCCAAAAGAGAGTGGTGCTTCCGAGCAAAATCTGATTGTTGCAAAATCA 2176
Qy 1981 tgggcagatgtagtataaacttgggtgcataaacaacacaaactaaagtctataaaaca tgggt 2040
Db 2177 TGGCAGATGTAGTAAAAACTTGGTGCATAAACAACAAACATAAAGTCAATAAAACATGGT 2236
Qy 2041 cctcaagagttcaatgaacaaaggcaagaagacctgctactctcaagaagcctgtgggc 2100
Db 2237 CCTCAAGGTCAATGAACAAAGGCAAGAAAGACCTGTACTCTCAAAAGAGCCGTGGGC 2296
Qy 2101 gaagttcacagttcaatttagtacaggccacgcgaactctcttgaccataataatagggt 2160
Db 2297 GAAGTTCACAGTCNAATTAGTACAGGCCACGCAAACTCTCTTGTACCATTAATAATAGGG 2356
Qy 2161 aaagctcactactgaaaaagtacatgtgctctgctgcgacccctacagagtgctcaacaattc 2220
Db 2357 AAAGCTCATACTGAAAAGTACATGTGCCCTGCTCGACCCCTACAGAGTGTCTCAACAACCTC 2416
Qy 2221 attttccacaaaaaatggacttttaagaaagatctttcaggaatagctgaaatttcaag 2280
Db 2417 ATTTCCAACCAAAAAATGGACTTTTAAGGAAGATCTTTTCAGGAATAGCTGAAATGTTCAAG 2476

Qy 2281 accccagtgaaagagcaaccgcagttgacaagcagcatgtcacatcgtctatttcaaatca 2340
Db 2477 ACCCCAGTGAAGGACCAACCCGAGTTGACAGACACATGTCACATCGCTATTTCAAATCA 2536
Qy 2341 gagaatttgcctggaaaaacagtttcaaggaaactgattcaggagaagaacctctgtctccc 2400
Db 2537 GAGAAATTTGCTTGGAAAACAGTTTCAAGGAACGTATTTCAGGAGAAGAACCTCTGTCTCCC 2596
Qy 2401 acctcaagaagtttggaggaaatgtgttcttcagtgacagaatgcagcaaaaacagcca 2460
Db 2597 ACCTCAGAGAGTTTGGAGGAAATGTGTCTTCTAGTCACAGATGTCAGCAAAACAGGCCA 2656
Qy 2461 tctgataaattcctctgcaagccctcccttaagagcgcgagtgatttagagaaaaatggaac 2520
Db 2657 TCTGATAAATCTCTGCNAAGCCCTCCCTTAAGACGGCAGTGTATTAGAGAAAATGGAAC 2716
Qy 2521 gtagcaaaaacgcccaggaaacacctacaaaatgacttctctgaggacaaaacttcagat 2580
Db 2717 GTAGCAAAAACGCCAGGAACACCTTACAAAATGACTTCTCTGGAGACAAAACCTTCAGAT 2776
Qy 2581 actgagacagagccttcaaaaacagtatccactgtataacaggtcaggaaggtctacaag 2640
Db 2777 ACTGAGACAGAGCCTTCAAAAACAGTATCCACTGTAAACAGGTCAGGAAGGTTCTACAG 2836
Qy 2641 ttcaggaaatacacagaagctacacctgtggaagtaagatgaagaaacaaaacacagaaatt 2700
Db 2837 TTCAGGAATATACAGAAGCTACCTCTGGAAAGTAAAGAGTGAAGAAACAANAATACAGAA 2896
Qy 2701 gttgagtgcatacctataaaaagggttcagaaggcaaacactactacaaagaagagagaaga 2760
Db 2897 GTTGAGTGCATCTCTAAAAGAGGTGAGAAGGCACACACTACTACAAACAAGGAGAGAAGA 2956
Qy 2761 gagatgaagaaatagaaagaccttttgagacataaagaaaattattgaattaaaagaa 2820
Db 2957 GAGATGAAGGAAATAGAAAGACCTTTTGAGACATATAAGGAAAATATTGAAATTAAGAA 3016
Qy 2821 aacgatgaaagatgaaagcaatgaagagatcaagaaacttggggcagaaatgtgcacca 2880
Db 3017 AACGATGAAAAGATGAAGCAATGAAGAGATCAAGAACTTGGGGGCAGAAATGTGCACCA 3076
Qy 2881 atgtctgaactgcagacacctcaagagcttgctgatacagaactcatgaagaacacgcca 2940
Db 3077 ATGTCTGACCTTGACAGAGCTCAAGAGCTTGGCTGTATACAGAACTCATGAAAGACACGCCA 3136
Qy 2941 cgtggccagaatctcctccaaacccaagatcatccaaagccacaaacagtgagaagagc 3000
Db 3137 CGTGGCCAGAAATCTCCTCCAAACCCCAAGATCATGCCAAGGCCAACAAAGAGTGAGAAAGGC 3196
Qy 3001 aaatcactaaaaatgccctgccagtcattacaacccagaaaccaaataaacaccccaacacac 3060
Db 3197 AAATCACTAAATGCCCTGCCAGTCAATTACAACCCAGAACCAATAAACCCCAACACAC 3256
Qy 3061 acaaaacaacagttgaaggcatcctctggggaagttagtgtagaagaagagctcctagca 3120
Db 3257 ACAAAACAACAGTTGAAGGCATCCTCTGGGAAAGTAGGTGTGAAAGAGAGAGTCTTAGCA 3316
Qy 3121 gtggcgaagttcacagagcgtcaggggagagccacgcacacacagacagagagccagcaga 3180
Db 3317 GTCGGCAAGTTTCACACGGACGTCAGGGGAGACCACGCACACGCACAGAGAGCCAGCAGGA 3376
Qy 3181 gatggcaagagcatcagaacgctttaagagtgctccaaagcagatccttgagccagcagcc 3240
Db 3377 GATGCAAGAGCATCAGAACGTTTAAAGGAGTCTCCAAAGCAGATCCTTGGACCCACAGCC 3436
Qy 3241 cgtgtaactggaatgaagaagtgcccaagaacgcgctaaggagagggcccagtcactagaa 3300
Db 3437 CGTGTAACTGGAATGAAGAGTGGCCAAAGACGCGCTAAGGAAGAGGCCCGCTCAGTAGAA 3496
Qy 3301 gacctggtgcttcaagagcctttccagacacagctccctctgaggaatcaatgact 3360
Db 3497 GACCTGGCTGCTTCAAAAGAGCTCTTCCAGACACACAGGTCCTCTGAGGAATCAATGACT 3556

QY 3361 gatgagaaactaccataaatagctgctgcaaatctccaccaccagaaatcagtggacactcca 3420
Db 3557 GATGAGAAAATACCAAAAATAGCTGCAAAATCTCCACCACCAAGAAATCAGTGGACACTCCA 3616
QY 3421 acaagcacaagaacaatggcctaagaagaagctctcagaaagcagatgtagaggaagaattc 3480
Db 3617 ACAGGCACAAAGCAATGGGCTAAGAGAAGTCTCAGGAAAGCAGATGTAGAGGAAGAAATTC 3676
QY 3481 ttagcactcaggaaactaacaccatcagcagggaagccatgctttacgcccacacagca 3540
Db 3677 TTAGCACTCAGGAAACTAACACCATCAGCAGGGAAGCCATGCTTACGCCCCAAACACGCA 3736
QY 3541 ggaagtgaatgagaagacattaaagcatttatyggaaactccagtcagaaactggacctg 3600
Db 3737 GGAGGTGATGAGAAAGACATTTAAAGCATTTATGGAACTCCAGTCGAGAACTGGACCTG 3796
QY 3601 cgagggaacttaccctgacgacaaaagacagctacagactcctaaggaaagagccagct 3660
Db 3797 GCAGGAACCTTTACCTGGCAGCAAAAGACAGCTACAGACTCCCTAAGGAAAGGCCCAGGCT 3856
QY 3661 ctagaagacctggctggctttaagagctcttccagactcctgggtcacaccgaggaattca 3720
Db 3857 CTAGAAGACTGGCTGGCTTTAAAGAGCTTCTCCAGACTCCTGGTCACACCGAGGAATTA 3916
QY 3721 gtggctgtctggtataaacactaaaaataccctgacacttccacagtcacagccagtgagc 3780
Db 3917 GTGGCTGTCTGGTAAACCACTAAAATACCTGCGACTCTCCAGCTCAGACCCAGTGGAC 3976
QY 3781 acccaacagacacaagacaagcccaagagaagtatcaggaagcagatgtagaggga 3840
Db 3977 ACCCCAAAGCACAAGCAAGCACCAGAGAGAGTATCAGGAAAGCAGATGTAGAGGGA 4036
QY 3841 gaactcttagctgtcagggaattctaaatgcccacagcaggcacaagccatgcacacgcctaaa 3900
Db 4037 GAACCTTAGCTGTCAGGAATCTAATGCCATCAGCAGGCAAGGCCATGCACACGCCCTAA 4096
QY 3901 ccatcagtaggtgaagagaaagacatcatcatatttgtgggaactccagtcagaaactg 3960
Db 4097 CCATCAGTAGTGAAGAGAAAGACATCATATTTGTGGGAATCCAGTCGAGAAACTG 4156
QY 3961 gacctgacagagaacttaaccgacagcagaagcggccaacaactcctaaggaagggcc 4020
Db 4157 GACCTGACAGAGAATTAACCGGACGAAGAGACGGCCCAAACTCCTAAGGAAAGAGGCC 4216
QY 4021 caggctctggaagacactgactggctttaagagctcttccagacccctgggtcatactgaa 4080
Db 4217 CAGGCTCTGGAAGACTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCATACTGAA 4276
QY 4081 gaagcagtaggtctgtggcaaaactactaaaaatgcccctggaatcttctccacagaatca 4140
Db 4277 GAAGCAGTGGCTGCTGGCAAAACTACTAAAATGCCCTGGCAATCTTCTCCACAGAAATCA 4336
QY 4141 cgagacaccccaacgaacgaagaaggcgcccaagacacctttggagaaaaggagcgtca 4200
Db 4337 GCAGACACCCCAACAGCACAGAGAGGAGGCCCAAGACACACTTTGGAGAAAAGGGACGTA 4396
QY 4201 cagaaggagctctcagccctgaagaagctcacacagacatcaggggaaaccacacacaca 4260
Db 4397 CAGAAGGAGCTCTCAGCCCTGAAGAAGCTCACACAGACATCAGGGGAAACCCACACACACA 4456
QY 4261 gataaagtaccagagggtgagatataaagcatcaacgcgttttagggaaactcgaataag 4320
Db 4457 GATAAAGTACCAGGAGGTGAGGATTAAGACATCAACGCCTTTAGGGAAACTGCAAAACAG 4516
QY 4321 aaactggacccagcagcaagttaactggttagcagagccaccccaaaactaaggaagaag 4380
Db 4517 AAATGGACCCAGACCAAGTGTAACTGGTAGACAGAGGCACCCAAAAACTAAGGAAAAG 4576
QY 4381 gcccaacccctagaagacctggctgtggaaagagctcttccagacacacagtatgcact 4440
Db 4577 GCCCAACCCCTAGAAGACCTGGCTGGCTGGAAGAGCTTCTCCAGACACCACTATGCAC 4636
QY 4441 gacaagcccaagcactcaagagaaaactaccacaaaatagcctgcagatcacaacacagacca 4500

Db 4637 GACAACCCACGACTCACGAGAAAACCTAGCAAAATAGCCTGCAGATCACAACCCAGACCCA 4696
QY 4501 gtggacacacccaacgaagtctccaagccacagtcacaagagagtcctcaggaagtgagcgtca 4560
Db 4697 GTGGACACACCAACAAGCTCCCAAGCCACACAGTCCAAGAGAAGTCTCAGGAAAGTGGACGTA 4756
QY 4561 gaagaagaattcttcgcactcaggaacaaacacacctcagcagcgaagcacaatgacaca 4620
Db 4757 GAAGAAGAAATCTTCCCACTCAGGAAACGACACCATCAGCAGGCAAAAGCCATGCACACA 4816
QY 4621 cccaaaccagcagtaagtgtgagaaaaacatactacgcatttatgggaactccagtycag 4680
Db 4817 CCCAAACCCAGCAGTAAGTGGTGTGAGAAAAACATCTAGGCATTTATGGGAATCTCCAGTGCAG 4876
QY 4681 aaactggacctcacagagaacttaactggcagaagaagcagcgtacaaaactcctaaggaa 4740
Db 4877 AAATGGACCTGACACAGAACTTAACCTGGCAGCAAGAGCGGTACAAACTCTCAANGGA 4936
QY 4741 aaggccagctctagaagacctggctgtttaagagctcttccagagctcttccagacagagtcac 4800
Db 4937 AAGGCCACGGCTCTAGAAGACCTGGCTGTAAAGAGCTCTTCCAGACACAGAGTCCAC 4996
QY 4801 actgaggaaatcaatgaactaacgataaaaactgccccaaagtgcctgcacaaatcttcaacaca 4860
Db 4997 ACTGAGGAATCAATGACTAACGATAAACTGCAAAAGTAGCCTGCAAAATCTTTCACAACCA 5056
QY 4861 gacctagacaaaaacccagacagctccaagcagcgtcacaagacatccctggggaagtg 4920
Db 5057 GACCTAGACAAAACCCAGCAAGCTCCAAGCGACGCTCAAGACATCCCTTGGGGAAGTG 5116
QY 4921 ggcgtgaagaagagagctcctcagttggcaagctcacacagacatcaggagagactaca 4980
Db 5117 GCGGTGAAAGAGAGCTCCTAGCAGTTGGCAAGCTCACACAGACATCAGAGAGACTACA 5176
QY 4981 cacacacacagagagccaaacagagagatgtgaagagcatgaaagcatttatgagctcca 5040
Db 5177 CACACACACACAGAGCCAAACAGAGATGTTAAGAGCATGAAAGCATTTATGGAGTCTCCA 5236
QY 5041 aagcagactcttagactcagcagcagcttaactggcagaagcagcagcagcagcagcagcct 5100
Db 5237 AAGCAGATCTTAGACTCAGCAGCAGCTTAACCTGGCAGCAAGAGCGCTGAGAACCTCT 5296
QY 5101 aagggaagctctgaagctcctgaagacctggcggcttcactcagagctcttccagacacca 5160
Db 5297 AAGGAAAAGTCTGAAGTCCCTGAAGACCTGGCGGCTTATCGAGCTCTTCCAGACACCA 5356
QY 5161 agtcacactaaggaatcaatgactaatgaaaaaactaccaaagctatcctcacagagcttca 5220
Db 5357 AGTCACACTAAGGAATCAATGACTAATGAAAACCTTACCAAAAGTATCCTTACAGAGCTTCA 5416
QY 5221 cagccagacctagtggaacaccccaacaagctccaagccacagcccaagagagctccagg 5280
Db 5417 CAGCCAGACCTAGTGGACACCCCAACAAGCTCCAAGCCACACGCCCAAGAGAGACTCAGG 5476
QY 5281 aaagcagacactgaagaagaatttttagcatttagaaacaaacgcatacagcaggcaaa 5340
Db 5477 AAAGCAGACACTGAAGAAGATTTTATAGCATTTAGGAAAACAAGGCCATCAGCAGGCAAA 5536
QY 5341 gccatgcacacaccccaaacccagcagtaggtgaaggaagaagacatacaacgctttttggga 5400
Db 5537 GCATGTCAACACCCCAACCCAGCAGTAGTGTAGTGAAGAAAGACATCAACACTGTTTTGGGA 5596
QY 5401 actccagtgcaaaaactggaccagccaggaatttaccctggcagcaatagacgctcacaa 5460
Db 5597 ACTCCAGTGCAGAAACTGGACAGCCAGGAAATTTACTTGGCAGCAANTAGACGGCTACAA 5656
QY 5461 actcgtgaaggaagcccgagctctcagaagaactgactggcttcagagagcttttccag 5520
Db 5657 ACTCGTAAGGAAAAGCCCGAGGCTCTAGAAGAACTGACTGGCTTCCAGAGAGCTTTTCCAG 5716
QY 5521 acaacatgcactgataacccccacagctgatgagaaactacccaaaaatactcttgcaaa 5580

Db 5717 ACACCATGCATGATAACCCACAGCTGATGAGAAAATACCAAAAAAATACTCTGCAA 5776
Qy tctccgcaatcagaccagcgagcaccccaacaaacaaacagcaacgcccacagagaagc 5640
Db 5777 TCTCGCAATCAGACCCAGCGACACCCCAACAACAACAGCAACGCGCCCAAGAGAGC 5836
Qy 5641 ctcaagaaagcagcgtatagagaagaaatttttagcatctcaggaactaacccatcagca 5700
Db 5837 CTCAGAAAGCAGACGTAGAGAGAGAAATTTTAGCATTTACAGAAACTAACCCATCAGCA 5896
Qy 5701 gcaaaagccatgcacacgcctaaagcagcagtaggtgaagaaagacatacaacattt 5760
Db 5897 GGCRAAGCCATGCACACGCCCTAAAGCAGCAGTAGGTGAAGAGAAAGACATCAACACATTT 5956
Qy 5761 gtgggactccagtggaaactggacctgtcgtagaaatttacctggcagcaagagacgg 5820
Db 5957 GTGGGACTCCAGTGGAGAACTGGACCTGCTAGGAATTTACCTGGCAGCAAGAGAGC 6016
Qy 5821 ccacaaactcctaagaaagccaaagcctctagaagatctggcttcaagagctc 5880
Db 6017 CCACAAATCTCTAAAGAAAGGCCAAGGCTCTAGAAGATCTGGCTGGCTTCAAAGAGCTC 6076
Qy 5881 ttccagacaccagctcacactgaggaatcaatgaccatgacaaaatcacagaagtatcc 5940
Db 6077 TTCAGACACCCAGGTCACACTGAGGAATCAATGACCGATGACAAATCACAGAAGTATCC 6136
Qy 5941 tgcgaatctcacacacagaccagctcaaaaccccaacaaagctccaaagcaacgactcaag 6000
Db 6137 TGCAAATCTCCACAAACCCAGCTCAAAACCCCAACAAGCTCCAAGCAAGACTCAAG 6196
Qy 6001 atatccttggggaagttaggtgtgaaagaagaggtcctaccagctggcgaactcacacag 6060
Db 6197 ATATCTTTGGGGAAGTAGGTGTGAAGAGAGGTCCTTACCAGTCGGCAAGCTCACACAG 6256
Qy 6061 acgtcagggaagacacacacacacagagagacagcagagagatggaagagcatcaaa 6120
Db 6257 ACGTCAGGGAAGACACACACACACAGAGAGAGCAGGAGATGGAAGAGCATCAAA 6316
Qy 6121 gcgtttaaggaattctgaaagcagatgctggaccacgaaactatggaactggagtgag 6180
Db 6317 GCGTTTAAGGAATCTGCAAAACAGATGCTGCACCCAGCAAACTATGGAACCTGGGATGGAG 6376
Qy 6191 aggtggccaaagacaccttaaggaagagcccaatcactagaagacctggccggcttcaaa 6240
Db 6377 AGGTGGCCCAAGAACCTTAAGGAAGAGGCCCCAATCACTAGAAGACCTGGCGGGCTTCAA 6436
Qy 6241 gagctcttcagacacagacacactgaggaatcaacaactgatgacaaactacaaa 6300
Db 6437 GAGCTCTTCCAGACACCCAGACACACTGAGGAATCAACAACTGATGACAAAACCTACCAAA 6496
Qy 6301 atagcctgcaaatctccaccaccagaaatcaatggacactccaaacagacagagggcg 6360
Db 6497 ATAGCTGCAAAATCTCCACCACCAAGATCANTGGACACTCCAACAAGCACAAGGAGCGG 6556
Qy 6361 cccaaaaacacttbygggaaagggatatagtggaagagctctcagccctgaagcagctc 6420
Db 6557 CCAAAAACACCTTTGGGGAAGGGGATATAGTGGAAAGAGCTCTCAGCCCTGAAGCAGCTC 6616
Qy 6421 acacagaccacacacagacaaagtaccagagatgaggaataaaggcatcaacgtgttc 6480
Db 6617 ACACAGACCAACACACAGACAAAGTACCAGGATAGGATAAAGGCATCAACGTGTTC 6676
Qy 6481 agggaaactgcaaaacagaaactggaccagcagcaagtgttaactggtagcaagaggcag 6540
Db 6677 AGGGAACCTGCMAAACAGAACTGGACCCAGCAGCAAGTGTAACTGGTAGCAAGAGGAG 6736
Qy 6541 ccaagaactcctaaggggaaagcccaaccccttagaagacttggcttggcttgaagagctc 6600
Db 6737 CCAAGAACTCTTAAGGGAAGAGCCCAACCCCTAGAAAGACTTTGGCTGGCTTGAAGAGCTC 6796
Qy 6601 ttccagacaccagtatgcactgacaagcccacgactcacgagaaaactacaaaatagcc 6660
Db 6797 TTCCAGACACCCAGTATGCATGACAAGCCCCAGACTCACGAGAAACCTACCAAAATAGCC 6856

Qy 6661 tgcagatctccacaaccagaccagtggtgtaccacaaatcttcaagccacagttccaa 6720
Db 6857 TGCAGATCTCCACACACAGAGCCAGTGGGTACCCCAACAATCTTCAAGCCACAGTCCAA 6916
Qy 6721 agaaatctcaggaagcagacagtagagaagaatccttagcactcaggaagaaacgaacaca 6780
Db 6917 AGAAGTCTCAGAAAGCAGACGCTAGAGGAAGAAATCTTTAGCACTCAGGAAAGCAACCA 6976
Qy 6781 tcaataggaagcttatgacacacccaaacacagcagaggtgatgaaagacatgaaa 6840
Db 6977 TCAGTAGGGAAGCTATGACACACCCAAACAGCAGGAGGTGATGAGAAACATGAAA 7036
Qy 6841 gcatttatggaaactccagtgacagaaattggacctgccaggaaatttacctggcagcaaa 6900
Db 7037 GCATTTATGGGAATCTCCAGTCAGAAATTTGGACCTGCCAGGAAATTTACCTGGCAGCAA 7096
Qy 6901 agatggccaaactcctaaggaagccaggtcctagaagacctggctggcttcaaa 6960
Db 7097 AGATGGCCACAACCTCTTAAGGAAAGGCCAGGCTCTAGAAAGAGCTGGCTGGCTTCAA 7156
Qy 6961 gagctctccagacaccaggaactgacaaagccacagctgatgaaactacaaaata 7020
Db 7157 GAGCTCTTCCAGACACAGGCACTGACAAGCCCCACGACTGATGAGAAACTTACCAATA 7216
Qy 7021 gctgcaaatctccacaaccagaccagtggaacccccagcagcaagcacaaagcaacggccc 7080
Db 7217 GCTGCAAAATCTCCACAACACAGACCCAGTGGACACCCCAAGCAAGCAACGGCCC 7276
Qy 7081 aagagaacctcaggaagcagacgttagaggagaatttttagcactcaggaagaaacgaaca 7140
Db 7277 AAGAGAAACCTCAGGAAAGCAGACGCTAGAGGAAGAAATTTTATAGCACTCAGGAAAGCA 7336
Qy 7141 ccacagcaggaagccatggacacccccaaacacagcagtaagtatgagaaaatc 7200
Db 7337 CCATCAGCAGGCAAGCCATGGACACCCCAAAACAGCAGTAAGTGTAGAGAAATATC 7396
Qy 7201 aacacatttggaaactccagtgacaaactggacctgctaggaatttacctggcagc 7260
Db 7397 AACACATTTGTGGAATCTCCAGTCAGAACTGGACCTGCTAGGAATTTTACCTGGCAGC 7456
Qy 7261 aagagacgcacagactccttaaggaagcgtgaggtcctagaggaacctggtgcttc 7320
Db 7457 AAGAGACACCCACACACTCTTAAGGAAAGGCTGAGGCTCTAGAGACCTGGTTGGCTTC 7516
Qy 7321 aaaaactctccagacaccaggtcacactaggaatcaatgactgatgacaaaatcaca 7380
Db 7517 AAAGAACTCTCCAGACACACAGGTCACTGAGGAATCAATGACTGATGACAAAATCACA 7576
Qy 7381 gaagtatcctgttaaattctccacagccagtgcatctcaaaacctcaagagctcccaagcaa 7440
Db 7577 GAAGTATCTCTGTAATCTCCACAGCCAGAGTCATTTCAAAACCTCAAGAAGCTCCAAGCAA 7636
Qy 7441 aggtcagaatcacccttggtagaagtgcacatgaaagagagccctagcagtcagcaag 7500
Db 7637 AGGCTCAAGATACCCCTTGTAAGATGGACATGAAAGAGAGAGCCCTAGCAGTCAGCAAG 7696
Qy 7501 ctcacagacacatcagggagactacgaaacacacacagacgacacagagagatagtaag 7560
Db 7697 CTCACCGGACATCAGGGGAGACTACGCAAAACACACAGAGCCCAACAGGATAGTAAG 7756
Qy 7561 agcataaagcgtttaagaggtctccaaagcagatccttgaccagcagcaagtgtaact 7620
Db 7757 AGCATCAAAAGCGTTTAAAGAGTCTCCAAAGCAGATCTTGGACCCAGCAGCAAGTGAAT 7816
Qy 7621 ggtagcaggaagcagctgagaactcgttaaggaagagcccgctgctctagaagacctggtt 7680
Db 7817 GGTAGCAGGAGCAGCTGAGAACTCTGTAAGGAAAGAGCCCGCTCTAGAGACCTGGTT 7876
Qy 7681 gacttcaagagctctctcagcaccaggtcacactgaagatcaatgactattacaaa 7740
Db 7877 GACTTCAAAGAGCTCTTCTCAGCACAGGTCACTGAAGAGTCAATGACTATTGACAAA 7936

QY	7741	aacacaaaaattccctgcgaatctccccaccagaactaacagacactgccacgacaca	7800
Db	7937	AACACAAAAATTCCTGCAATCTCCCCACCAGAACTTAACAGACACTGCCACGAGCACA	7996
QY	7801	aagagatgcccagaacagctccagagaaagtaaaagaggagctctcagcagttgag	7860
Db	7997	AAGAGATGCCCCCAAGACACGCTCCAGAGAAAGTAAAAGAGGAGGCTCTCAGCAGTTGAG	8056
QY	7861	aggctcacgcaaacatcagggcgaagcacacacacacacaaagaaacagcaagcggtgat	7920
Db	8057	AGSCTCACGCAACATCAGGGCAAGCACACACACACACAAAGAACAGCAACGCGTGAT	8116
QY	7921	gagggcatcaaaqatgtgaagcaacbtgcacaaagaaagaaacccacagtagaaggaa	7980
Db	8117	GAGGGCATCAAGTATTGAAGCAACCTGCAAAAGAAAGAAACCAACCCAGTAGAAGAGGAA	8176
QY	7981	cccagcagagagagccaagagcacctaaggaaaaggcccaacccctggaagacctggcc	8040
Db	8177	CCAGCAGGAGAGGCCAAGAGCACCTTAAGGAAAAGGCCCAACCCCTGGAAGACCTTGGCC	8236
QY	8041	ggcttcacagagctctctgaacatcaggttcacactcaggaaatcactgactgctggcaaa	8100
Db	8237	GGCTTCACAGAGCTCTGAAACATCAGGTTCACACTCAGGAATCACTGACTGCTGGCANA	8296
QY	8101	gccactaaaaatccctgcgaatctccccactagaagtgttagacaccacagcaagaca	8160
Db	8297	GCCACTAAAATACCTTCCGAATCTCCCCACTAGAAAGTGTAGACACCACAGCAAGCACA	8356
QY	8161	aagagcatctcagagacagtgacagaaggtacaagttaaagaagagccttcagcagtc	8220
Db	8357	AAGAGGCATCTCAGGACACGCTGCGAGAGGTACAAGTAAAAGAAAGGCTTCAGCAGTC	8416
QY	8221	aagttcacacaaatcaggggaaaccaggtgcagacaaagaaacacagcaggtgaagt	8280
Db	8417	AAGTTCACACAAACATCAGGGGAAACCAGGATGCAGACANAGAACACAGCAGGTGAAGT	8476
QY	8281	aaaggcatcaaaqatgtgaaggaaatgtcaaaacagacacccggctccagcagcaagtga	8340
Db	8477	AAAGGCATCAAAACATTTGAAGGAATCTGCAAAACACACACACCCGGCTCCAGCAGCAAGTGA	8536
QY	8341	actggcagcagagagcgccaagagcaccagggaaagtgcccaagccatagaagacctta	8400
Db	8537	ACTGGCAGCAGAGACGGCCAGAGCACCCAGGGAAAGTGGCCCAAGCCATAGAAGACCTA	8596
QY	8401	gctggctcaaaagccagcagcaggttcacactgaagaatcaatgactgatgacaaaacc	8460
Db	8597	GCTGGCTTCAAAACCCAGCAGCAGGTCACTGAAGATCAATGACTGATGACAAAACC	8656
QY	8461	actaaaatccctgcgaatctaccagaaactagaagacacccgcaacagctcaaaagaga	8520
Db	8657	ACTAAAATACCCCTGCAAAATCATCCACGAACCTAGAAGACACCCCAACAGCTCAAAAGA	8716
QY	8521	cgsgccagagacagtgccacgaaagtagaagtgaaggagagctgttagcagttggcaag	8580
Db	8717	CGGCCCAGACAGTCCCCAGAAAGTGAAGTGAAGAGGAGAGCTGTTACAGTTGGCAAG	8776
QY	8581	ctcacacaaacctcaggggagacacgcacacgcgcaaaagagcggtagtgagggcaaa	8640
Db	8777	CTCACAAAAACCTCAGGGGAGACACCGCACACCGCAAAAGAGCGCGTAGTGAGGGCAAA	8836
QY	8641	ggcacgaagcatttaagcaacctgcaagcggaagcagtggaagcagaagaatgtaattggc	8700
Db	8837	GGCAGAAAAGCATTTAAGCAACCTGCAAAAGCGGAACCTGGAGCGCAGAAAGATGTAATTGC	8896
QY	8701	agcagagacagccaagagcacctaaggaaaaggcccaacccctggaagacctggccagc	8760
Db	8897	AGCNGAGACAGCCAGACGACCTTAGGAAAAGGCCCAACCCCTGGAAGACCTTGGCCAGC	8956
QY	8761	ttccaagagctctctcaaacaccagggccacactgaggaactggcaaatggtgctgctgat	8820
Db	8957	TTCCAAAGAGCTCTCTCAAAACACCAGGCCACACTGAGGAACCTGGCAAAATGGTGTCTGAT	9016
QY	8821	agctttacaagcgtccaaagcaaacacactgacagtggaagaccttaaaaaatatccaga	8880

Db	9017	AGCTTTACAGCGCTCCAAAGCAAAACACCTGACAGTGGAAACCTCTTAAATAATATCCAGA	9076
QY	8881	aaagtctctcgggccctctaaagttagaacccgttggaagcgttgtaagcaccagagacct	8940
Db	9077	AGAGTTCTTCGGGGCCCTTAAAGTAGAACCCGTGGGAGACGTGGTAAGCACACAGAGACCT	9136
QY	8941	gtaaatcacaaagcaaaagcaacttcctctccccactgccttccaagagggaggt	9000
Db	9137	GTAATAATCAAAAGCAAAAGCAACACTTCCTTCGCCCACTGCCCTTCAAGAGGGGAGGT	9196
QY	9001	ggcaagatggaaagcgtcacgggaacaaagagcgtgcgtgcatagcagcaccagagaaa	9060
Db	9197	GGCAAAAGATGGAAAGCTCACGGGAACCAAGAGCTGCGTGCATGCGAGCACCAGAGGAA	9256
QY	9061	attgtgagagagctgcccagcagcaagacagagcagaggttgctcccgagggcagaagcaaa	9120
Db	9257	ATTGTGGAGAGGCTCCAGCCAGCAAGACAGAGGTGTCTCCAGGGGNAGAGCAAA	9316
QY	9121	tcatacgaacccgtggtcatcatgaagagaagtttgaggacttctgcataaaagaattgaa	9180
Db	9317	TCATCCGAACCCGTGCTCATCATGAAGAGAAGTTTCAGGACTTTCGCAAAAACAATTGAA	9376
QY	9181	cctgcggaagagctgaacagcagcatgaaaaaccaacaaagagagaaacaaattacaa	9240
Db	9377	CCTGCGAAAGAGCTGAACAGCAACGACATGAAACCAACAAAGAGGGAACAACAATTACAA	9436
QY	9241	gactcgttcctgaaataaagggaatatccctgcctccagacgccaagaataagactgag	9300
Db	9437	GACTCGTCCCTGAAATTAAGGGAATATCCCTCGCTCCAGAGCCCAAGATAAGACTGAG	9496
QY	9301	gcagaacagcaataaactgaggtctttgttatagcagaagaatagaaaataaacagaaaa	9360
Db	9497	GCAGAACACACAATAACTGAGGTCTTTGTATTATAGCAGAGAAGATAGAAATAAACAGAAAT	9556
QY	9361	gaaaagaagcccatgaagacctccccagagatggacatttcagaattccagatgtgagcc	9420
Db	9557	GAAAGAAGCCCATGAAGACCTTCCCAGAGATGGACATTCAGAAATCCAGATGATGGAGCC	9616
QY	9421	cggaaccccatcactagagacaaaagtctactgagaacaaaggctgcttgaggtctgctaga	9480
Db	9617	CGGAAACCCATACCTAGACACAAAGTCACTGAGACAAAGAGTGTCTTGAGGTCTGTAGA	9676
QY	9481	cagaaatgagagctcccagcctaaggttggcagagagcggagggcagaagagtcggaag	9540
Db	9677	CAGAAATGAGAGCTCCAGGCTTAAGGTGGCAGAGAGAGCGGAGGCAGAAAGAGTGGCAAG	9736
QY	9541	gttctcatgcagaatcagaaaagggaaggaagcaggaatttcagactccatgtgcctg	9600
Db	9737	GTCTCATGCAGAAATCAGAAAGGGAAGGAGAGCAGGAAATTCAGACTCCATGTGCCTG	9796
QY	9601	agatcaagaaagacaaaaggccagcctgcagcaagcactttggagagcaaatctgtgcag	9660
Db	9797	AGATCAAGAAGACANAAAGCCAGCTGCAGCAAGCAGCTTTGGAGAGCAAAATCTGTGCAG	9856
QY	9661	agagtcaacgcgaggtgtcaagaggtgtgcagaaaaataccaaagagggctgaggaacaatgtg	9720
Db	9857	AGAGTAACCGGAGTGTCAAGAGGTGTGAGAAAATCCAAAAGAAAGGCTGAGGACAATGTG	9916
QY	9721	tggttcaagaaaaatacaaccaggaagtcattagggacagtgaaagata	9766
Db	9917	TGTGTCAAGAAAAATAACAACCCAGAAAGTCATAGGGCAGTGAAGATA	9962

RESULT 2

AX336950

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AX336950 Sequence 12515 bp DNA linear PAT 09-JAN-2002

AX336950 from Patent WO0194629.

AX336950 .

AX336950.1 GI:18127669

human.

Homo sapiens

Db 2057 GTGGCAACCTGCCTTCAAAAGAGAGTGTCTATCAGCGCAAGTCAACATGATATTTACAG 2116
QY 1921 atgatatgttccaaagaagaagtgtgtcttcggaagcaaatctgattgttgcaaatca 1980
Db 2117 ATGATATGTTCCAAAGAGAAGTGTGTCTCGAAGCAAAATCTGATTTGTGCAAAATCA 2176
QY 1981 tgggcagatgtagtagtaaaacttgggtcacaacaaacaaactaaagtcatataaacttgg 2040
Db 2177 TGGGCAGATGTAGTAAAACTTGGTGTCAAAACAAACAACTAAAGTCATATAAACATGGT 2236
QY 2041 cctcaaggtcattgacaaaaggcaaaagacacctgtctactccaagaagcgtgtgggc 2100
Db 2237 CCTCAAGGTCAATGAACAAAGGCAAAAGGCAAGAACCTGTCTACTCCAAAGAAAGCCTGTGGCC 2296
QY 2101 gaagtccacagtcatttagtacaggccacgcgaacctctcctgttaccataataataggg 2160
Db 2297 GAAGTTCACAGTCAATTTAGTACAGGCCACGCNAACCTCTCTTGTACCATATAATATAGGG 2356
QY 2161 aaagctcactgaaaaagtacatgtgcctgtctcgacctacagagtgtctcaacaacttc 2220
Db 2357 AAAGCTCATACTGAAAAAGTACATGTGCCTGTCTGACCCCTACAGAGTGTCTCAACAACTTC 2416
QY 2221 atttccaaacaaaagtgaactttaaggagaagatctttcagggaatagctgaaatgttcaag 2280
Db 2417 ATTTCCAACCCAAAATGGACTTTTAAGGAAGATCTTTTCAGGAATAGCTGAAATGTTTCAAG 2476
QY 2281 accccagtgaagagcaacccgcagtttgacaagcacatgtccacatcgctatttcaaatcca 2340
Db 2477 ACCCCAGTGAAGGCAACCCGCAGTTGACAGCACATGTACATCTGCTATTTCAAAATTTCA 2536
QY 2341 gagaaattgcttgaaaaacagtttcaaggaactgattcaggagaagaacctctgtccccc 2400
Db 2537 GAGAAATTGCTTGGAAAAACAGTTTCAAGGAACGTGATTCAGGAGAAGCACTCTGCTCCCC 2596
QY 2401 acctcagaagtttggaggaatgtgttcttcagtgacagaatgacagcaaaacagcca 2460
Db 2597 ACCTCAGAGAGTTTGGAGGAATGTGTCTTCAGTGCACGAATGCAGCAAAACAGCCA 2656
QY 2461 tctgataaatgtctctgaagccctcccttaagacgagctgatttagagaaaatggaac 2520
Db 2657 TCTGATAAATGTCTTCGAAGCCCTCCCTTAAGACGSCAGTGTATTAGAGAAATGGAAAC 2716
QY 2521 tagcaaaaaccccggaacacctcaaaaatgactctctctggagacaaaactctcagat 2580
Db 2717 GTAGCAAAAACGCCAGGAACACCTACAAAATGACTTCTCTGGAGACAAAACCTTCAGAT 2776
QY 2581 actgagacagagccttcaaaaacagtatccactgtaaacaggtcaggaaggtctacagag 2640
Db 2777 ACTGAGACAGAGCCTTCAAAAACAGTATCCACTGTAAACAGGTTCAGGAAGGTCTACAGAG 2836
QY 2641 ttcaggaatatcacagaagctacctgtggaagtgaagtgaaagtaagaacaaatcacagaaat 2700
Db 2837 TTCAGGAATATACAGAAGCTACCTGTGGAAAGTAAAGATGAAGTAAGAAACAATACAGAAAT 2896
QY 2701 gttgagctactctaaaaagaggtcagaaggcaaacactactacaacaaaggagagaagga 2760
Db 2897 GTTGTGTGATCTCAAAAAGAGGTCAAGAGGCAAGGCAACACTACTACAACAAAGGAGAGAAGGA 2956
QY 2761 gagatgaaggaatagaagaccttttgagacatatagaagaaaaatttgaattaaagaa 2820
Db 2957 GAGATGAAGGAATAAGAAAGACCTTTTGTGACATATTAAGGAATAATTTGAATTTAAAGAA 3016
QY 2821 aacga tgaagaatgaagcaatgaagaga tcaagaaccttggggcagagaaatgtgcacca 2880
Db 3017 AACGATGAAGAATGAACCAATGAAGAGATCAAGAACTTGGGGCAGAGAAATGTGCACCA 3076
QY 2881 atgtctgacctgacagacctcaagagcttgcctgtatcacagaactcatgaagacagcca 2940
Db 3077 ATGCTGTGACCTGACAGACCTCAAGAGCTTGCCTGTATACAGAACTCATGAAGACACGGCA 3136
QY 2941 cgtggccagaatctcctccaaaccaaagatcatgtccaagggcaccaaagagtgaagaagc 3000

Db 3137 CGTGGCAGAAATCTCTCCAAACCCCAAGATCATGTCCCAAGGCACCAAGAGTGTGAAAGGC 3196
QY 3001 aaaaactcaataaattgccttgcagtcattacaaccagaaccaaataaacaccccaacacac 3060
Db 3197 AAATCACTAAATGGCTTGGCAGTCAATTACAACCCAGAACCAATAAACACCCCAACACAC 3256
QY 3061 aaaaaaacagltgaagtcaccccttgggaaaagttaggtgtgaaaagaagagctccttagca 3120
Db 3257 ACAAAAACAACAGTTGAAGGCATCCCTGGGAAAGTAGGTGTGAAAGAAGAGCTCCTTAGCA 3316
QY 3121 gtccgcaagtccaacgagcgtcaggggagacacacgacacgacacagagagcagcagga 3180
Db 3317 GTCCGCAAGTTACACGGACGTCAGGGGAGACACACGACACAGAGAGCCAGCAGGA 3376
QY 3181 gatggcaagagatcagaacgtttaaggagctcccaagcagatccttgagaccagcagcc 3240
Db 3377 GATGGCAAGACATCAGAACGTTTAAGGAGTCTCCAAAGCAGATCTCTGGACCCAGCAGCC 3436
QY 3241 cgtgtaacttggaaatgaagaatgagcaagccctaaagaaagagggcccgctcacttagaa 3300
Db 3437 CGTGTAACTGGAATGAAGAAGTGGCCAAGACGCTTAAGGAAGAGGCCCTAGCAGTCACTAGAA 3496
QY 3301 gacctggtcgttcaaaagagctcttccagacacacaggtccctctcaggaatacaatgact 3360
Db 3497 GACCTTGGCTGGCTTCAAGAGAGCTTTTCCAGACACACAGGTCCTCTGAGGAATCAATGACT 3556
QY 3361 gatgagaaactaccaaaatagcctgcaaatctccaccaccagaaatcagtggaactcca 3420
Db 3557 GATGAGAAAACCTACCAAAATAGCCTGCAAAATCTCCACCACCAGAAATCAGTGGACACTCCA 3616
QY 3421 acaagcacaagaacaaatggccttaagagaagttctcaggaaaagcagatgtagagggaagattc 3480
Db 3617 ACAAGCACAAAGCAATGGCTTAAGAGAAGTCTCAGGAAAGCAGATGTAGAGGAAGAATTC 3676
QY 3481 ttgacctcaggaanaactaacaccatcagcagggaaagccatgcttaccgcccacacagca 3540
Db 3677 TTAGCACTCAGGAAACTAACACCATCAGCAGGGAAAGCCATGCTTACGCCCAAAACAGCA 3736
QY 3541 ggagggtgagagaagacatataagcatattatgggaactccagtgacagaaacttggacctg 3600
Db 3737 GGAGGTGATGAGAAAGACATTTAAAGCATTTATGGGAATCTCAGTGCAGAAACTGGACCTG 3796
QY 3601 gaagaaacttacctggcggcaaaagacagctacagactccttaaggaanaagggcccggt 3660
Db 3797 GCAGGAACCTTACCTGGCAGCAAAAGACAGCTACAGACTCCTAAGGAAAGGCCAGGCT 3856
QY 3661 ctagaagacctgctggtgtttaaagagctcttccagactccttggtcacaccaggaattta 3720
Db 3857 CTAGAAAGACCTGGCTGGCTTTTAAAGAGCTCTTCCAGACTCCTTGGTCACACCGGGAATTA 3916
QY 3721 gtggctgtgtgtaaaaccactaaaa taccctgcgactctccacagtcagacccagtgagac 3780
Db 3917 GTGGCTGTGTGTTAAACCACTTAAATACCTTGCAGCTCTCCACAGTCAGACCCAGTGGAC 3976
QY 3781 acccaacaagaacaaagcaacgaccccaagagaagtatcaggaagaagcagatgtagaggga 3840
Db 3977 ACCCCAACAAGCACAAAGCAACGACCCCAAGAGAAGTATCAGGAAAGCAGATGTGTAGAGGA 4036
QY 3841 gaactcttagcgtcaggaatctaatgtccatcagcagggcaaaagccatgcacagccctaaa 3900
Db 4037 GAACCTTTAGCTGTGAGGAATCTAATGGCCATCAGCAGGCAAAAGCCATGCACAGCCCTAAA 4096
QY 3901 ccactagtaggtgaagagaagaacatcatatttctgtgggaactcctcagtcagagaactg 3960
Db 4097 CCATCAGTAGGTGAAGAGAAAGACATCATATTTTGTGGGAACCTCCAGTCAGAAACTG 4156
QY 3961 gacctgacagagaacttaaccggcagcagaagagcggccacaactccttaagaagaaggcc 4020
Db 4157 GACCTGACAGAGAACCTTAACCGCGCAGCAAGAGACGCGCCACAAACTCCTTAAGGAAGAGCC 4216
QY 4021 caggtcttgaagacactgactggtctttaagagactcttccagacccctgtcatactgaa 4080
Db 4217 CAGGCTCTGGAAGACCTGACTGGCTTTTAAAGAGCTCTTCCAGACCCCTGTCTATCTGAA 4276

D	b	6437	 GAGCTCTTCCAGACACACACACACACTGAGGAATCAACAACCTGATGACAAAACCTACCCAAA	6496
Q	y	6301	atagcctgcgaatctccacacacagaatcaatggacactccaaacaaacacaaagcagagcg	6360
D	b	6497	ATAGCTGCAAAATCTCCACCACAGANTCAATGGACATCTCAACAAGCACANGAGGCGG	6356
Q	y	6361	cccaaacacctttggggaaggatagtggaagagctctcagccctgaagcagctc	6420
D	b	6557	CCCAAAACACCTTTGGGGAAGGGATATAGTGAAGAGCTCTCAGCCCTGAACAGCTC	6616
Q	y	6421	acacagaccacacacagacaaagttaccaggagataggataaaggcatcaacgtgttc	6480
D	b	6617	ACACAGCACACACACACAAAGTACCAGGAGATGAGGATAAAGGCATCAACGTGTC	6676
Q	y	6481	agggaactctcaaacacagaaactggaccacagcaagtgaactggtacgaagagcag	6540
D	b	6677	AGGAAACTGCAAAACAGAAACTGGACCCAGCAGCAAGTGAACCTGCTGAGCAAGGCGAG	6736
Q	y	6541	ccaagaactcctaagggaagcccaaacctctagaagacttggctggcttgaagagctc	6600
D	b	6737	CCRAGAACTCCTAAGGGAAAGCCCAACCCCTAGAAAGACTTGGCTGGCTTGAAGAGCTC	6796
Q	y	6601	ttccagacaccagta tgcactgacaagcccagctcacgagaaaactaccaaatagcc	6660
D	b	6797	TTCCAGACACACCTATGCACTACAAAGCCAGCTACGAGAAAACTACCCAAATAGCC	6856
Q	y	6661	tgagatctccacaacacagaccagtggtgtaccccaacactctcaagccaagccaag	6720
D	b	6857	TGCAAGATCTCCACACACAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAGTCCAAAG	6916
Q	y	6721	agaagtctcaggaaagcagctagaggaagaatccttagcactcaggaaacgaacacca	6780
D	b	6917	AGAAGTCTCAGGAAACACAGCTAGNGAAGAAATCCTTAGCACTCAGGAACGAAACCA	6976
Q	y	6781	tcsgtgggaaagctatggacacaccccaacacagcagggtgatgtagaagaagacatgaaa	6840
D	b	6977	TCAGTAGGAAAGACTATGGACACACCCAAACAGCAGCAGGCTGATGAGAAAGACATGAAA	7036
Q	y	6841	gcattatggaaactccagtgcagaaattggactgcaggaaaatttacctggcagcaaa	6900
D	b	7037	GCATTTATGGGAACTCCAGTGCAGAAATTTGGACCTGCCAGGAAATTTACCTGGCAGCAAA	7096
Q	y	6901	agatggccaaaactcctaaggaaaaggccaggctctagaagacctggctggctccaaa	6960
D	b	7097	AGATGGCCACAAATCCTTAAGGAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTCAAA	7156
Q	y	6961	gagctcttcagacacacaggcactgacaagcccagctgatgagaaaaactaccaaaata	7020
D	b	7157	GAGCTCTTCCAGACACACAGGCATGCAAGGCCAGCTGATGAGAAAACTACCCAAATA	7216
Q	y	7021	gcctgcgaatctccacaacacagaccagtggaaccccgacgaagcaacagcaagccc	7080
D	b	7217	GCCTGCAAAATCTCCACACACAGCCAGTGGACACCCAGCAAGCAAAAGCAACGGCCC	7276
Q	y	7081	aagagaaactcaggaaagcagctagaggaagaatttttagcactcaggaaacgaaca	7140
D	b	7277	AAGAGAACTCAGGAAACAGACAGCTAGAGGGAAGAAATTTTAGCACTCAGGAACGAAACA	7336
Q	y	7141	ccatcagcaggcaagccattgacaccccccaaacccagcagtaagtgtagaaaaata tctc	7200
D	b	7337	CCATCAGCAGGCAAAAGCCATGGACACCCCAACCCAGCAGTAACTGTATGAGAAAAATATC	7396
Q	y	7201	aacacattgtggaactccagtgcagaaacttgacctgtctaggaaaatttacctggcagc	7260
D	b	7397	AACACATTTGTGGAAACTCCATGTGAGAAACTTGGACCTTGTAGGAAATTTACCTGGCAGC	7456
Q	y	7261	aagagacagccacagactcctaaggaaaagcctgagctcttagaggacctggttgcttc	7320
D	b	7457	AAGAGACGCCACAGACTCCTAAGGNAAGGCTGAGGCTCTTAGAGGACCTGGTTGGCTTC	7516
Q	y	7321	aaagaactcttcagacacacaggtcacactgaggaatcaatgactgtagcaaaaatcaca	7380

Db	7517	AAAGAACTCTTTCCAGACACACAGGTCTACACTGAGGAATCAATGACTGATTCACAAAATCACA	7517
Qy	7381	gaagtatccctgtaaatctccacagccagagtcattcaaaacctcaagagctcccaagcaa	7440
Db	7577	GAAGTATCCTGTAAATCTCCACAGCCAGAGTCAATTCAAAACCTCAAGAAGCTCCAAAGCA	7636
Qy	7441	aggtctaagataccctctggtagaagtggacatagaagaagagcccttagcagctcagcaag	7500
Db	7637	AGGCTCAAGATACCCCTGGTGAAGTGGACATGMAAAGAGAGCCCCCTAGCAGTCAGCAAG	7696
Qy	7501	ctcacacggacatacaggggagactacgcaaacacacacagagccaaacggaggatagtaag	7560
Db	7697	CTCACGGGACATCAGGGGAGACTAGCAGAAACACACACAGAGCCAAACAGGAGATAGTAAG	7756
Qy	7561	agcatcaaacggtttaagaggtctccaaagcagatccttgaccacagcagcaagtgtaact	7620
Db	7757	AGCATCAAGCGTTTAAAGGAGTCTCCAAAGCAGATCCTGGACCCAGCAGCAAGTGTAACT	7816
Qy	7621	ggtagcagggagcagctggaactctgaagaaaagcccggtgctcttagaagacctggtt	7680
Db	7817	GGTAGCAGGAGGCAGCTGAACACTCGTAGGAAAAGGCCCGTGCTTAGAAGACCTGGTT	7876
Qy	7681	gacttcaaaagagctctctcagcaaccaggttcacactgaagagtcactgactattgacaaa	7740
Db	7877	GACTTCAAGAGGTCTTCTCAGCACAGGTTCACACTGAAGAGTCAATGACTATTTGACAAA	7936
Qy	7741	aacacaaaattccctgcgaattctccccaccagaaactaacagacacitgccacagacaca	7800
Db	7937	ANCAAAAATTCCTTGCAANTCTCCCCACACAGAACTATACAGACACTGCCACAGACACA	7996
Qy	7801	aagagatgccccaagacacgtcccgaggaagaagttaaagaagagagctctctcagcagttgag	7860
Db	7997	AAGAGATGCCCAAGACACGTCCCGAGAAAGTAAAGAGAGGAGCTCTCAGCAGTTGAG	8056
Qy	7861	aggtctacgcaaacatacagggcaaaagcacacacacacacacaaagaacccagcagcggtag	7920
Db	8057	AGGCTCAGCAAAACATCAGGGCAAAAGCACACACACACAAAAGAACACAGCAAGCGGTGAT	8116
Qy	7921	gagggcatcaagattatgaagcaacgtcgaagaagaacccaaacccagtagtaagaagaaa	7980
Db	8117	GAGGGCATCAAGTATTGTGAAGCAACGTGCANAGAGAAACCAACCCAGTATGNAAGAGGNA	8176
Qy	7981	cccagcaggaagaaggccaagagcacctaaaggaaaaggcccaaccccttggaagacctggcc	8040
Db	8177	CCACAGCAGGAAGAGGCCAACAGCACCTAAGGAAAAGGCCCAACCCCTGGAAGACCTGGCC	8236
Qy	8041	ggcttcacagagctctctgaaacatacaggttcacactcaggaatacactgactgctggaaca	8100
Db	8237	GGCTTCACAGAGTCTCTGAACATCAGGTCTCACACTCAGGAATCACTGACTGCTGGCANA	8296
Qy	8101	gcacataaataccctgcgaattctccccacactagaagtggtagacaccacacagcagacaca	8160
Db	8297	GCCACTAAAATACCTTCGGNANTCTCCCCACCTAGAAGTGGTAGACACCACAGCAAGACACA	8356
Qy	8161	aagaggcatcttcaggacacgtgtgcagaaggtacaaagtaaagaagagaccttcagcagctc	8220
Db	8357	AAGAGGCATCTCAGGACACGTGTGCAGAGAGGTACAAGTAAAGAAGAGCCCTTCAGCAGTC	8416
Qy	8221	aagttcacacaaacatacaggggaaacccacggtgcagacaaagaacaccagcgagtgaaagat	8280
Db	8417	AAGTTTACACAAAACATCAGGGGAAAACCCCGGATGCGACACAAGAAGAACACAGCAGGTCAAGAT	8476
Qy	8281	aaaggcatcaaacattgaaggaattctgaaacacagacacggctccagcagcagtgta	8340
Db	8477	AAAGGCATCAAGCATTTGAAGGAATCTGCANAACAGACACCGGGCTCCAGCAGCAAGTGTA	8536
Qy	8341	actggcagcaggaagcggccaagagcacccagggaagtgcccaagcca tagaagacctta	8400
Db	8537	ACTGSCAGCAGGAGACGGCCACAGACACCCAGGAAAGTGCCTCAAGACCATAGAAGACCTA	8596
Qy	8401	gctggcttcaaaagaccacagcagcaggtcacactgaagaatacactgactgtatgcacaaacc	8460
Db	8597	GCTGGCTTAAAGACCAGCAGCAGGTACACTGAAGAAATCAATGACTGATTGACAAAACC	8656

Qy 8461 actaaataccctgcacaaatcatccacagaaactagaaagacacaccgcaacagctccaagaga 8520
|||||
Db 8657 ACTAAATACCTGTCAAATCATCCAGAACTAGAAAGACACCCGCAACAGCTCAAGAGA 8716
Qy 8521 cggcccaagacacgtgcccagaaatagaaagtgaaggaggagctgttagcagtgccaag 8580
|||||
Db 8717 CGGCCAGGACACGTGCCCAGAAATAGAACTGAAGTGAAGGAGGAGCTGTTCAGATTTGCCAAG 8776
Qy 8581 ctcacaaaaactcaggggagaccacgcacaccgacaaagagccgttaggtgagggcaaa 8640
|||||
Db 8777 CTCACAAAACTCAGGGAGACACCGCACACCGACAAAGAGCCCGTAGGTAGGGCCAAA 8836
Qy 8641 ggcagaaagcatttaagcaacctgcacaaagggaacgtggacgcgaagaagtgaattggc 8700
|||||
Db 8837 GGCACGAAAGCATTTAAGCAACCTCAAAGCGGAACGTGGACGCAGAAATGTAATTGCG 8896
Qy 8701 agcaggagacagcaagagcacctaaggaagaaagcccaacccctggaagacctggccagc 8760
|||||
Db 8897 AGCAGGAGACGCCAAGAGCACCTTAAGGAAAGGCCCAACCCCTGGAAGACCTGGCCACG 8956
Qy 8761 ttccaagagctctccaacacacagccacactgaggaactggcgaatggctgctgat 8820
|||||
Db 8957 TTCCAAGAGCTCTCAACACACGACGCCACACTGAGGAACTGGCAAAATGGTCTCTGAT 9016
Qy 8821 agctttacaagcgtctccaaagcaaacacctgacagtggaaaaacctctaaaaatatccaga 8880
|||||
Db 9017 AGCTTTACAAGCGCTCCAAGCAACACACTGACACTGGAAACCTCTAAAAATATCCAGA 9076
Qy 8881 agagttctcgggcccctaaagtagaacccgtggagacgtggtgaagcaccagagacct 8940
|||||
Db 9077 AGAGTCTTCGGGGCCCTTAAAGTAGAACCCGCTGGGAGACGTGGTAAGCACCCAGACCCCT 9136
Qy 8941 gtaaatcacaaagcaaaagcaaacacttccctgccccacactgccttcaaaagggaggt 9000
|||||
Db 9137 GTAAATCACAAGCAAAAGCAACACTTCCCTGCCCCCACCCTGCTCCCTTCAAGAGGGAGGT 9196
Qy 9001 ggcaagatgaagcgtgcacgggaaccagagagcgtgcgtcatccgacccagcagagaa 9060
|||||
Db 9197 GCAAAAGATGGAAGCGTCACGGGAACCAAGAGCGCTGCGTGCATGCCAGCACAGAGAA 9256
Qy 9061 attgaggaggagctgcagcagcaagaagcagagaggtgtgtccagggcaagagcaaa 9120
|||||
Db 9257 ATTGTGGAGAGCTCCACGACGACGAAGAGCAGAGGGGTGCTCCAGGCGCAAGAGCAAA 9316
Qy 9121 tcattccgaacccgtgttcattcatatgaagaaagattgaggaactctgcataaagaattgaa 9180
|||||
Db 9317 TCATCCGAACCCGTGCTCATATGAAGAGAACTTTGAGGACTTCTGCAGAAAGAAATTGAA 9376
Qy 9181 cctgcggaagagctgaacagcaacgacatgaaaaaccaaaccaagagagaacacaaattacaa 9240
|||||
Db 9377 CCTGCGGAAGAGCTGAACAGCAACGACATGAAACCAACAAAGAGGAACACAAATTACAA 9436
Qy 9241 gactcgttccttgaaaaataagggaatatccctgcgtccagcgcagacgccaagataagactgag 9300
|||||
Db 9437 GACTCGGTCCCTGAAAATTAAGGGAATATCCCTGCGCTCCAGACGCCAAGATAAAGACTGAG 9496
Qy 9301 gcagacagcaaaataactgaggtctttgtattagcagaaagaatagaaataaacagaaat 9360
|||||
Db 9497 CGAGAACAGCAATTAACCTGAGGCTTTGTATTAGCAGAAAGAAATAGAAATAAACAGAAAT 9556
Qy 9361 gaaaaagagcccatgaagacacctccccagagatggacattcagaatccagatgatggacc 9420
|||||
Db 9557 GAAAAAGAGCCCATGAAGACCTCCCCAGAGATGGACATTCAGAATCCAGATCATGGACC 9616
Qy 9421 cggaaacccatctctagagacaaagtcaactagagacaaagaagtgctgaggtcctctaga 9480
|||||
Db 9617 CGAAACCCATACCTTAGAGACAAAGTCACTGAGAACAAAAGGTGCTTGAGGTCTGCTAGA 9676
Qy 9481 cagaatgagactccagccctaaagttggcagagagagcgagggcgacagaaagtcgcaag 9540
|||||
Db 9677 CAGAAATGAGACTCCAGCCCTAAGTGTGCAGAGGAGCGGGAGGCAGAAAGTCCGGAAG 9736

Qy 9541 gtctctatgcagaatcagaagggaaagagagcaggagaaattcacactccatgtgcctg 9600
|||||
Db 9737 GTTCTCATCGAATCAGAAAGGGAAGGAGAGCAAGAAATTCAGACTCCATGTGCCTG 9796
Qy 9601 agatcaagaaagacaaaaagccagcctgcagcaagcactttggagagacaaatctgtgcag 9660
|||||
Db 9797 AGATCAAGAAAGACAAAAAGCCAGCCTGCAGCAAGCACCTTGGAGAGCAAAATCTGTGCAG 9856
Qy 9661 agagtcaacgcggagtgctcaagaggtgctgagaaaaatccaaaaagagcctgaggaacaatgtg 9720
|||||
Db 9857 AGAGTAACCGGAGTGTCAAGAGGTGTGCAGAAAAATCCAAAAGAGCTGAGGACAATGTG 9916
Qy 9721 tgtgtcaagaaataacaaccagaaagtcatagggacagtgaaagata 9766
|||||
Db 9917 TGTGTCAAGAAAAATAACAACCAAGACTCATAGGACAGTGAAGATA 9962

RESULT 3
HSMK167
LOCUS
DEFINITION H.sapiens mki67a mRNA (long type) for antigen of monoclonal antibody Ki-67.
ACCESSION X65550
VERSION X65550.1 GI:415818
KEYWORDS antigen; monoclonal antibody.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12515)
AUTHORS Gerdes Fors, J.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1992) J. Gerdes, Forschungsinstitut Borstel, Div. Molecular Immunology, Parkallee 22, 2061 Borstel, FRG
REMARK sequence revised by author 13-JUL-93 and 08-OCT-93
AUTHORS 2 (bases 1 to 12515)
Schluter, C., Duchrow, M., Wohlenberg, C., Becker, M.H., Key, G., Flad, H.D. and Gerdes, J.
TITLE The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins

J. Cell Biol. 123 (3), 513-522 (1993).

94043435 Location/Qualifiers
1..12515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10q25-qter"
/cell_line="IM9"
/clone_lib="lambda gt11"
1..107
/number=1
108..288
/number=2
197..9967
/gene="mki67"
197..9967
/gene="mki67"
/codon_start=1
/product="antigen of the monoclonal antibody Ki-67"
/protein_id="CAA46519.1"
/db_xref="GI:415819"
/translation="MWPTRLVTIKRSGVDGPHFLSLSTCLFGRGICEDIRIOLPVV SKOCHKIEHQEALIHNFSSNTPTQVNGSVIDEPVRLKHGVITFIIDRSFRYENSL QNGKSTEFPRKIREQPARVRSRSFSDPDEKADSKAYSKEGKVSQPVQHIK KVEDSTADDSKDSVAQGTNTVHSEHAGRNRAADPTSGDFKEISSVKLYSRGEL NSVPTQCIDNSKNESFPWKLYESVKKELDVYKSNENLOYCRKSGLTDTATSKES ADLQGETOLLVSRKSRPKSGGSHAVAPASPELDQNKCKGRDVSQVTPSKAVG ASFPLYEPKMKPTVOYSOQONSPOKHKNKDYTTGRRSVNLKSGCFKAGDKTLTP RKLSTRNRTPAKVEDAASATKPNELSSKTRGSIPTDVEVLPTETEHNEPPLTLWT QVERIKQDSLSUKPLTGTAGCMCSGLPGLSSVDINNFGDSINSESEGLPLKRRRVSF GGHLPRLFDENLPPNTPILKRGEAPTKRKSLVMHTPPVLKRIKIQPOPSGQESGSE

I HVEVKAQSLVISPAPSPRKTPVASDQRRRCKTAPASSSKSQTEVPRGGERVATC
LQKRVISRSQHDILQMICSKRRSGASEANLIVAKSWADVVLKGAKQTQTQTVIKHGFPQ
RSMKQRPRPQKPVGVSHQSFGTHANSPICTIIIGKAHTEKVKHVPARPVRLNPF
ISNQKMDGSLGIAEMFKTPVKEOQLTSCIIAISNSENLLGKFOGQTSGBEPL
LPTSESFGNVTFSAONAAKQPSDKCSASPLLRQOCI RENGNAVKTPTNTYKWSLET
KQSDTEPESKTVSTVNSGRSTEFNIOQLPVESSEETNTEIVECILKROKATLL
QORREGEKEIERPFETKENIELKENDENKMRKSTWGQRCAPMSDLTDLKSLPD
TELMKOTAKGONLLOQDHAKAPKSEKIKTKMPCOSLOPEPINTPTHTKQOLKASLG
KVGKKEELAVLAKFTTSGTTHREPAAGDKSIRTFKESPKQILDPAARVTGMKKW
PRTPKEEASLEDLAGKELFQTPPGSESMDEKTKIACKSPPESDVDTPTSTKOW
PKRSLRKADVEEFALRLKLTSPAGKAMLT PKPAGGDEKDIKAFMGTPVKQDLDACTL
PGSKROLOTPKPKAQALEDLAGKELFQTPGHTHEELVAAGKTKT IKCDSPQSDPVDTTP
TSTKQPKKSIKRAKADVEGELLACNLMPAGKAMHTPKPSVGEEDKIIIFVGTVPVOKL
DUTENUTGSKRRPQTPKEBAQALEDLTGFKELFQTPGHTHEEVAAGKTKTKMPCESP
ESADTPTSTRRQPTPLEKRDVQKELSAKLTQTSGETTHDKVPGEDKSI NAFRE
TAKQILDPAASVTGSKRHKPTKEKAQPLEDLAGWKELFQTPVTKDPTTHETTKIAC
RSOPDVPPTSSKPSKSLRKVDVEEFFALRKRTPSAGKAMHTPKPAVSEKNIY
AFMGTPVKQILDLENLTGSKRLQTPKKAQALEDLAGKELFQTRGHTHEESWINDKT
AKVACKSPDDDKNPASSKRLKTSIGKVGVEELLAVGKLTQTSGETTHHTHTPTG
DKSKMAFQSPKQILDSASLTGSKRQLRTPKGSSEVPEDLAGTLELFTQTPSHTKES
MNEKTKVYSRASQPDLDVPTSSKQPKRSLRKADTEEEFLAFRQTPSAGKAMHT
PKPVGGEEDKINTFLGTPVKQILDQPNLPGSNRRLOTRKEKAQLELTFGRLEFQTP
CTDNPTADKTKTKILCKSQSPADPTNTKQRPKRSILKADVEEFALFRLKLPSPA
GRAMHTPKAAGEEKDINTFVGTVPVEKLDLGNLPGSKRRPQTPREKAKALEDLGKFK
ELFQTPGHTHEESWINDKITEVCKSPQDPVKTPTSSKQRLKISLGKVGKVEVLPGV
KLUTQTSKTKTQTHRETAGDGKSIKAPKESAKQMLDPA NTYGTGMRWPRTPKBEAOSLE
DLAGKELFQTPDHTHEESTDDTKTKIACKSPPPSMDTPTSTRRPQTPGLKRDIVE
ELSAKQLQTPHTHDKVPGEDKGINVRETAKQILDPAASVTGSKRQRPTPKGRAQ
LEDLAGKELFQTPVTKDPTTHETTKIACRSOPDPVGTPTIPKPSKSLRKADV
EESLALRKRTPSVKAMDTPKPAGGDEKMAFMGTVPVKQLDLGNLPGSKRWQTP
KKAQALEDLAGKELFQTPGDKPTTDEKTKIACKSPQDPVDTPTASTKQRPKNL
RKADVEEFLALRKRTPSAGKAMDTPKPAVSDEKNTINTEVPVKQLDLGNLPGSKR
QTPKEKAEALDLVGKELFQTPGHTHEESMTDDKITEVCKSQPSQSPESFQTSRSKQ
RLKIPLVKDMKEEPLAVSKLRTSGTTHETPTGDSKSIKAPKESPKQILDPAAS
VTGSRQLRTRKEKALEDLVDEKELFSAPGHTESMTIDKNTIKTPCKSPPELTD
ATSTKPCRTREPKVKEELSAVERLTQTSQSTHTHKEPASDGEIGKVLKORAKKP
NPVEEPPSRRRPRAPKEKAQLEDLAGTSELSETSGHTOESLTGAKATKIPCESPPLE
VDTTASTKRLHRLRVQKVQVKEEPSAVKFTQTSGETTHDADPEAGEDGKIKALKE
KQTPAPASVTSRRPRAPRAESAQIEDLAGFKDPAAGHTEESMTDDKTKTIPCKSS
PELEDTASTSKRRPTRAQKVEVELLAVGKLTQTSGETTHDKEPVEGEGADSKFK
OPAKRNVAEDVIGSRROPRAPKEKAQPLEDLASFOELSOTPGHTEELANGADSF
APKQTPDSCKPLKISRRVLRAKVPVGVSTROPVKSQKSNLSPLPKRGGGK
DCSVGTKRLKMPAPEEIVBELPASKKORVAPRAKGSSEPVVIMKRSURTSAKRIE
PAELNLSKMPKNEEHLKQDSVPENKGLSLRRRODKTEAEQQITEVEFLARIEIN
RNEKPKMTSPMDTQPNDDGARKPIPRDKVTENRCLRSARONESSQPKVAEESGQ
KSAKVLNQKQKGAGNSDMLRSRKTQSOPAASTLESKSVQRVTRSVKRCENPK
KAEDNVCKKII TRSHRSEDI"
289. .367
/gene="mk167"
/number=3
368. .483
/gene="mk167"
/number=4
484. .550
/gene="mk167"
/number=5
561. .596
/gene="mk167"
/number=6
597. .1676
/gene="mk167"
/note="partially excluded by splicing"
/number=7
1677. .1852
/gene="mk167"
/number=8
1853. .2165
/gene="mk167"
/number=9
2166. .2284
/gene="mk167"
/number=10
2285. .2456

/gene="mk167"
/number=11
2457. .2612
/gene="mk167"
/number=12
2613. .9457
/gene="mk167"
/number=13
9458. .9901
/gene="mk167"
/number=14
9902. .12493
/number=15
polyA_signal
10529. .10534
polyA_signal
11237. .11242
polyA_signal
12468. .12473
polyA_site
12494
BASE COUNT 4166 a 3048 c 2928 g 2373 t
ORIGIN
Query Match 99.5%; Score 9715; DB 9; Length 12515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9765; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 atgtgccccagagagccctgttactatcaaaaggagcggggtcgacggTccccacttt 60
|||||
Db 197 ATGTGGCCACGAGAGCGCTGGTTACTATCAAAAGGAGCGGGTTCGACGGTCCCCACTTT 256
QY 61 cccctgagcctcagcactcgttctgttggagggggtattgaaatgtgcataccgtatccag 120
|||||
Db 257 CCCTGAGCCTCAGCACCTGTTGTTGGAAAGGGGTATTGAATGTGCATCCGTATCCAG 316
QY 121 ctctcgttgttcgaaaaacaacattgcaaatgtgaaatccatccatgagcaggaggaattta 180
|||||
Db 317 CTTCCTGTTGTGCAAAAACAACATTGCAAAATTCGAAATCCATGACGAGGAGGAATATTA 376
QY 181 caLaatttcagtctccacaacatccacaacagtaaaatgggtctgttattgatgagcctgta 240
|||||
Db 377 CATAATTTCAAGTTCCACAAATCCACACACAAAGTAATGGGTCTGTTATTGATGAGCCTGTA 436
QY 241 cggctaaaaacatcggagatgtaatactattcattgatcgttcccttcagggtatgaaaaTgaa 300
|||||
Db 437 CGGCTAAAACATGGAGATGTAATAACTATTATTGATCGTTCGTTCCGTACAGTATGAAAATGAA 496
QY 301 agtcttcagaatggaaggagtcactgaactgaatttccagaagaaatacgtgaacaggagcca 360
|||||
Db 497 AGTCTTCAGAAATGGAAGGAGTAGTCAACTGAAATTTCCAGAAAAAATACGTGAACAGGAGCCA 556
QY 361 gcacgtcgtgtctcaagatctagctctctcttgacccctgatgagaagctcaagattcc 420
|||||
Db 557 GCACGTCGTGTCACAGATCTAGCTTCTCTTCGACCCCTCATGAGAAAGCTCAAGATTCC 616
QY 421 aagcctattcaaaaatcactgaagaaaaagtttcaggaaaaatcctcagggtacatatcaag 480
|||||
Db 617 AAGGCTATTTCAAAAATCACTGAAGGAAAAAGTTTCAGGAAATCCTCAGGTACATATCAAG 676
QY 481 aatgtcaaaagaacagtagtaccgcagatgactcaaaagacagtgattgtctcagggaacaact 540
|||||
Db 677 AATGTCAAAAGAAGCAGTACCGCAGATGACTCAAAAGACAGTGTGTGCTAGGGAACAACT 736
QY 541 aatgttcattcctcagaacaatcgtggagcgttaatggcagaaatcagctgatccccattct 600
|||||
Db 737 AATGTTTCATTCCTCAGAACATGCTGGAGCTTAATGGCAGAAATGCAGCTGATCCATTCT 796
QY 601 ggggaatttaaaagaatttccagcgtttaaattagtgagccgttataggagaattgaagtct 660
|||||
Db 797 GGGGATTTTAAAGAAATTTCCAGCGTTAAATTTAGTGAGCGGTTTATGAGAGAAATTTGAAGTCT 856
QY 661 gtccccactacacaatgtcttgacaatagcaaaaaaaatgaatctccctctttggaagctt 720
|||||
Db 857 GTTCCCACTACACAATGTCTTTGACAAATAGCAAAAAAATGAATCTCCCTTTTGGAGCTT 916

QY	721	tatgagtcagtggaagaaagagttggtgtataaatacaaaaagaaaaatgtctactacagtat	780
DB	917	tatcagtcagtggaagaaagagttggtgtataaatacaaaaagaaaaatgtctactacagtat	976
QY	781	tgtagaaaatcttgattacaactgattacgcaacagagaagaaagtgtctgtagtggttta	840
DB	977	tgtagaaaatcttgattacaactgattacgcaacagagaagaaagtgtctgtagtggttta	1036
QY	841	cagggggagaccacaactgttggtctcgcgtgaagtcacaagaccacaaaatctggtggagcggc	900
DB	1037	cagggggagaccacaactgttggtctcgcgtgaagtcacaagaccacaaaatctggtggagcggc	1096
QY	901	cagcctgtggcagagcctgttcaactgaacagagcttgaccagaacaaggggaagga	960
DB	1097	cagcctgtggcagagcctgttcaactgaacagagcttgaccagaacaaggggaagga	1156
QY	961	agagacgtggagctgtgtttcagactccacgaaggtgtggtggcgcagctttctctctat	1020
DB	1157	agagacgtggagctgtgtttcagactccacgaaggtgtggtggcgcagctttctctctat	1216
QY	1021	gagccggctaaaatgaagaccacctgtacaatatccacagcaacaaaattctccacaaaa	1080
DB	1217	gagccggctaaaatgaagaccacctgtacaatatccacagcaacaaaattctccacaaaa	1276
QY	1081	cataagaacaagacctgtatactactggtagaagagaatctgtgaatctgggtataaagt	1140
DB	1277	cataagaacaagacctgtatactactggtagaagagaatctgtgaatctgggtataaagt	1336
QY	1141	gaaggtctcaaggtgtgtgataaaactcttactccaggaagcttccaactagaatactga	1200
DB	1337	gaaggtctcaaggtgtgtgataaaactcttactccaggaagcttccaactagaatactga	1396
QY	1201	acaccagctaaagttgaagatcagctgactctgccactaagccagacaaaatctctcttc	1260
DB	1397	acaccagctaaagttgaagatcagctgactctgccactaagccagacaaaatctctcttc	1456
QY	1261	aaaccagaggaagtattctcagatgtgaagttctgcctacggaactgaattcac	1320
DB	1457	aaaccagaggaagtattctcagatgtgaagttctgcctacggaactgaattcac	1516
QY	1321	aatgagccatttttaactctggctcactcaagttagaggaagatccaaaaggaattcc	1380
DB	1517	aatgagccatttttaactctggctcactcaagttagaggaagatccaaaaggaattcc	1576
QY	1381	ctcagcaagcctgagaaaattgggcaactacagctgacagatgtgctctbggttacctggt	1440
DB	1577	ctcagcaagcctgagaaaattgggcaactacagctgacagatgtgctctbggttacctggt	1636
QY	1441	cttagttcagttgatatcaacaactttggtattccattaatgagatgaggaataacct	1500
DB	1637	cttagttcagttgatatcaacaactttggtattccattaatgagatgaggaataacct	1696
QY	1501	ttagaaagaagcggtgtcttcttgggtggccacttaagacctgaactatttgaagaaac	1560
DB	1697	ttagaaagaagcggtgtcttcttgggtggccacttaagacctgaactatttgaagaaac	1756
QY	1561	tgcctctctaatacgccctcctcaaaaggggagaagccccaacaaaagaagttctctggtta	1620
DB	1757	tgcctctctaatacgccctcctcaaaaggggagaagccccaacaaaagaagttctctggtta	1816
QY	1621	atgcacactccacctgtctcgaagaaatcatcaggaagacagcctcaacatcaggaaa	1680
DB	1817	atgcacactccacctgtctcgaagaaatcatcaggaagacagcctcaacatcaggaaa	1876
QY	1681	caagagtcagggttcagaaaatccatgtggaagtgaaggcacaagaacttggttataagccct	1740
DB	1877	caagagtcagggttcagaaaatccatgtggaagtgaaggcacaagaacttggttataagccct	1936
QY	1741	ccagctctactagtcctaggaaaactccagttgccagtgatcaacgcctgaggtctctgcaaa	1800
DB	1937	ccagctctactagtcctaggaaaactccagttgccagtgatcaacgcctgaggtctctgcaaa	1996

Qy	1801	acagccctgtcttccagcagcaaaatctcaagacagaggttctctaagagagagagagaaaga	1861
Db	1997	ACAGCCCTGCTTCCAGCAGCAAAATCTCAGACAGAGGTTCTTAAGACAGAGGAGAAAGA	2056
Qy	1861	gtgccaacctgccttcaaaagagagtgctatcagccgaagtcaacatgatatttaccag	1920
Db	2057	GTGGCAACCTGCCTTCAAAGAGAGTGTCTATCACCAGTAGTCAACATGATATTTTACAG	2116
Qy	1921	atgatattgtccaaaaagaagaagtgtgtcttcggaagcaaatctgtattgttgcaaaaatca	1980
Db	2117	ATGATATGTTCCAAAAAGAAAGTGTGTCTCGGAAGCAAAATCTGATTGTTGCAAAATCA	2176
Qy	1981	tgggcagatgtatgaaaaacttggtgcaaaaaacaacataaagtcatataaacatcgtt	2040
Db	2177	TGGSCAGATGTAGTAAAACTTGTGCAAAACAAACACAAACTAAAGTCTATAAAAACATGGT	2236
Qy	2041	cttcaaaagtccaatgaacaaaaggcaaaagaagacctgtactccaaagaagcctgtgggc	2100
Db	2237	CCTCAAAAGTCAATGAACAAAAGGCAAGAAGCCTGTCTCTCAAGAAGCCTGTGGGC	2296
Qy	2101	gaagttcacagtcgaatttagtcacagccacgcgaacactctccttgtaccaataataaggy	2160
Db	2297	GAAGTTCACAGTCAAAATTTAGTACAGCCACGCAAACTCTCCTGTACCATAAATAAGGG	2356
Qy	2161	aaagtcatactgaaaaagatcatgtgtcctgtcgacctacagagtgctcaacaacttc	2220
Db	2357	AAAGCTCATACTGAAAAGATACATGTGCTGTCTGCACCTACAGAGTGTCAACAACCTC	2416
Qy	2221	atttccaaacaaaaatggaactttaaggagaatcttttcaggaaatagctgaaatttcaag	2280
Db	2417	ATTTCACACCAAAAATGACATTTAAGGAAGATCTTTCAGGAATAGCTGGAATGTTCAG	2476
Qy	2281	acccagtgaaaggagcaaccgcagttgacaagacacatgtcacatcgctatttcaaatcca	2340
Db	2477	ACCCAGTGAAGGAGCACCGCAGTTGACAGCACATGTACATCGCTATTTTCAAAATCA	2536
Qy	2341	gagaattgtcttgaaaaacagtttcaaggaaactgattcaggagagaacacctctgtctccc	2400
Db	2537	GAGAAATTTGCTTGGAAAACAGTTTCAAGGAACATGATTACAGGAGAGAACCCTCTGCTCCC	2596
Qy	2401	acctcagagagtttggagaaatgtgttcttcaatgcacagaatgcagcaaaacagcca	2460
Db	2597	ACCTCAGAGAGTTTTGGAGGAAATGTGTTCTAGCTGCACAGAATGCAGCAAAAACGCCA	2656
Qy	2461	tctgataaatgtctctcgaagccctccccttaagcggcagtgattagagaaaaatggaaac	2520
Db	2657	TCTGATAAATGCTCTCGACGCCCTCCCTTAAGCGGCAGTGTATTAGAGAAAATGGAAAC	2716
Qy	2521	gtagcaaaaacgcccaggaaacacctacaaaatgaactctctggagacaaaacttcagat	2580
Db	2717	GTAGCAAAAACGCCAGGAACACCTACAAAATGACTTCTCTGGAGACAAAAAATTCAGAT	2776
Qy	2581	actgagacagagccttcaaaaacagtaataccactgtaaacaggtcagggaaggtctacagag	2640
Db	2777	ACTGAGACAGAGCCTTCAAAAACAGTATCCACTGTAAACAGGTCTCAGGAAGGTCTACAGAG	2836
Qy	2641	ttcagaaatacacagagctactctgtgaaaagtgaagatgaagaaacaaatcacagaatt	2700
Db	2837	TTCAGGAAATATACAGAAGCTTACTCTGTGAAAGTAGAGTGAAGAAACAAATACAGAAATT	2896
Qy	2701	gttgagtgcactctaaaaagaggtcagaaggcaacactactacaacaaagagagagaagga	2760
Db	2897	GTTGAGTGCATCTTAAAAAGAGGTGAGAGGCAACACTACTACAACAAAAGGAGAGAAGGA	2956
Qy	2761	gagatgaaggaaatagaagaacctttgagacataaaggaaaaatattgaattaaaagaa	2820
Db	2957	GAGATGAAGGAAATAGAAAACGCTTTTGAGACATATAAGGAAAATATTGAAATTAAGAA	3016
Qy	2821	aacgatgaaaagatgaacgaatgaagaatcaagaacttggggcgagaatgtgcacca	2880
Db	3017	AACGATGAAAAGATGAAGCAATGAAGAGATCAAGAATCTGGGGCAGAAATGTGCACCA	3076
Qy	2881	atgtctgaacctgcacacacctcaagagcttgcctgatacagaactcataaagaacacgcca	2940

Db 3077 ATGTCTGACCTGACAGACCTCAAGAGCTTGCCTGTATACAGAACTCATGAAGACACGGCA 3136
QY cgtggccagaaatctctctccaaacccaagatcatgccaaggcaccacaagagtggagaaggc 3000
Db 3137 CGTGGCCAGAAATCTCTCTCCAAACCCAAAGATCATGTCGAAGGCACCAAGAGAGTGAGAAAGGC 3196
QY aaaaactcaaaaatgcctgcagcttcataaccagaaacaataaaacacccccaacac 3060
Db 3197 AAAATCACTAAAATGCCCTGCCAGTCAATTACAACCCAGAACCAATATAACACCCCAACACAC 3256
QY aaaaaacaagttgaagcatccctggggaagtagtgtgaagaagagctccttagca 3120
Db 3257 ACAAACAACAGTTGAAGGATCCCTGGGGAAGTAGTGTGAAGAAGAGCTCTCTAGCA 3316
QY gtcggcaagttcacacgcagtcgaggagagaccacgcacacgcacagagagccagcagga 3180
Db 3317 GTCGGCAAGTTCACACGCGCTCAGGGGACACCCACGCACACAGAGAGCCAGCAGGA 3376
QY gatggcaagagcatcagaaacgtttaaggagctctccaaagcagatccttggaccacagcc 3240
Db 3377 GATGGCAAGAGCATCAGAACGTTTAAAGAGTCTCCAAAGCAGATCCTTGGACCCAGCAGCC 3436
QY cgtgtaactggaatgaagaagtgcacaagacccctaaggaagagggcccaagtcactagaa 3300
Db 3437 CGTGTAACTGGAATGAAGAAGTGGCCAAAGAACGCCTAAGGAAGAGGCCAGTCACTAGAA 3496
QY gacctggctgcttcaaaagagctcttcagacaccaggtccctctgaggaatcaatgact 3360
Db 3497 GACCTGGCTGCTTCAAAGAGCTTCTCCACACACCAGGTGCTCCCTCTGAGGAATCAATGACT 3556
QY gatgagaaactaccaaaaatagcctgcaaatctccaccaccagaaatcagtggaactcca 3420
Db 3557 GATGAGAAAATACCAAAAATAGCTGCAAAATCTCCAAATCTCCACACAGAAATCAGTGGACACTCCA 3616
QY acaagcaaaagcaatggcctaagaaagtcctcagaaagcagatgtagaggagaattc 3480
Db 3617 ACAAGCAAAAGCAATGGCCCTAGAGAAGTCTCAGGAAGCAGATGTAGAGGAAGAAATTC 3676
QY ttgacctcaggaaaactaacaccatcagcagggaagccatgcttacgcaccaaacccagca 3540
Db 3677 TTAGCACTCAGGAACTAACCACCATCAGCAGGGAAGCCATGTCTACGCCCAACACGCA 3736
QY ggagggtagagaaagacataaagcatttatgggaactccagtcagaaaaactggacctg 3600
Db 3737 GGAGGGTGATGAGAAAGACATTAAGGATTTATGGGAATCTCCAGTGCAGAACTGGACCTG 3796
QY gcaggaaacttacctggcagcaaaagacagctacagactccttaagaaaagggccaggct 3660
Db 3797 GCAGGAACTTTACTGGCAGCAAAACACAGCTACAGACTCTTAAGGAAGGCCCAAGGCT 3856
QY ctagaagacctggctggctttaagagctcttcagactccttggtcacaccggaggaaatita 3720
Db 3857 CTAGAAGACCTGGCTGTAAAGAGCTCTTCCAGACTCCTTGGTCACACCCGAGGAATTA 3916
QY gtgctgtcgtgtaaaaccactaaaaataccctgcgaactctccacagtcagaaccagtcggac 3780
Db 3917 GTGGCTGCTGGTAAAACCACTAAAATACCCCTGGCACTCTCCACAGTCAGACCCAGTGGAC 3976
QY accccaacagcaaaagcaacgaccaccaagagaagtatcaggaagcagatgtagaggga 3840
Db 3977 ACCCCAAACAGCAAAAGCAACGACCCAGAGAAAGTATCAGGAAGACAGATGTAGAGGGA 4036
QY gaactcttagctgcagggaattctaattgcatcagcaggcaaaagccatgcacacgcctaaa 3900
Db 4037 GAACTCTTAGCTGTCAGGAATCTTAATGCCATCAGCAGGCAAAAGCCATGCACAGCCCTAAA 4096
QY ccatcagtaggtgaagagaaagacatacaatatttgggaaactccagtgcagaactcg 3960
Db 4097 CCATCAGTAGGTGAAGAGAAAGACATCATATTTTGTGGGAATCTCCAGTGCAGAAACTG 4156
QY gacctgacagagaacttaaccggcagcagagacggccacaactccttaaggaagggcc 4020

Db 4157 GACCTGACAGAGAACTTTAACGGCAGCAAGAGAGCGGCACAAAATCTCTTAAGGAAGAGGCC 4216
QY caggctctggaagacctgactggctttaagagagctctctccagaccctctggcataactgaa 4080
Db 4217 CAGGCTCTGGAAGACCTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCATACTGAA 4276
QY gaagcagtggtctgtggcaaaactactaaaatgcccttcgcgaattcttctccaccagaaatca 4140
Db 4277 GAAGCAGTGGCTGCTGGCAAAACTACTATAATGCCCTGCGAATCTTCTCCACCAGAAATCA 4336
QY gcagacacccccacaagcacaaagcagccccaagacacctttgggagaaaaagggacgta 4200
Db 4337 GCAGACACCCCAACAAGCACAAAGAGCGCCCAAGACACCTTTGGGAGAAAAGGGACGTA 4396
QY cagaagagagctctcagccctgaagagctcacacagacatacaggggaaacccacacacaca 4260
Db 4397 CAGAAAGAGCTCTCAGCCCTGAAGAGCTCACACAGACATCAGGGGAAACCCACACACACA 4456
QY gataaangtaccagaggtgaggataaaagcatcaaacgcgttttaggggaaactgcacaaacag 4320
Db 4457 GATAAAGTACCAGAGGTGAGGATAAAAGCATCAACGCGTTTAGGGAAACTGCAAAACAG 4516
QY aaactggaccacacagcaagtcgaagcttagcaagagggcaccacaaactaaggaaaaag 4380
Db 4517 AAATGGACCCACGAGCAAGTGAACCTGGTAGCAAGAGGCCCAAAAATTAAGGAAAAG 4576
QY gcccaacccctagaagacctggctggctggaagagctctctccagacaccagtatgcact 4440
Db 4577 GCCCAACCCCTACAGACCTGGCTGGCTGGAAAGAGCTCTTCCAGACACCAGTATGGACT 4636
QY gacaagccacgactcacagagaaaaactaccaaataagcctgcagatcacaaacccagaccca 4500
Db 4637 GACAAGCCACGACTCACGAGAAAATATACCAAAATAGCCTGCAGATCACAACACAGACCCA 4696
QY gtgacacacccaagaagctccaagccacagtcgaagaagagctctcaggaagaagtggacgta 4560
Db 4697 GTGGACACACCAACAGCTCCAAGCCACAGTCCAAGAGAAGTCTCAAGAAAGTGGACGTA 4756
QY gaagagaaattctcgactcaggaagaaacccaaccaaccagagcaagcattgcacaca 4620
Db 4757 GAAGAAATTCCTTCGCACTCAGAAACGAAACACCATCAGCAGGCAAGCCATGCACACA 4816
QY cccaaacccagcagtagtggtggagaaaaacatctacgcatcttatgggaaactccagtcgag 4680
Db 4817 CCCAAACACGACATAGTGGTGAGAAAAACATCTACGCATTTATGGGAACCTCCAGTGCAG 4876
QY aaactggacctgacagagaacttaactggcagcaagagacgctacaaaactcctaaggaa 4740
Db 4877 AAATCGACCTGACAGAGAACTTAACCTGGCAGCAAGAGAGCGGCTACAAATCTCTAAGGAA 4936
QY aaggcccgaggtctagaagacctggctggctttaagagagctcttccagacacagaggtcac 4800
Db 4937 AAGGCCCAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACAGGCTCAC 4996
QY actgaggaatcaatgactaacgataaaaactgccaagtgagcctgagcctgcacaaacttcacaacca 4860
Db 4997 ACTGAGGAATCAATGACTAACGATAAAACTGCCAAAAGTAGCCTGCAAAATCTTTCACAACCA 5056
QY gacctagacaaaaacccaagaagctccaagcgagcgctcaagacatccctcggggaaagtg 4920
Db 5057 GACCTAGACAAAANCCAGCAAGCTCAAGCGACGGCTCAAGACATCCCTTGGGGAAGTG 5116
QY ggcgtgaaagaagagctccttagcagttggcaagctcacacagacatcagagagagactaca 4980
Db 5117 GCGGTGAAAGAGAGGCTCTTAGCAGTTGGCAAGCTCACACAGACATCAGGAGAGACTACA 5176
QY cacacacacagagccacacagagagatgtaagagcatgaagaatcttatggaggtctcca 5040
Db 5177 CACACACACAGAGCCAAACAGGAGATGGTAAAGAGATGAAGAGCATTAAGAGTCTCCA 5236
QY aagcagatcttagactcagcagcaagcttaactggcagcaagagcgagctgcgaactcct 5100
Db 5237 AAGCAGATCTTAGACTCAGCAGCAAGTCTAACTGGCAGCAAGAGCGAGCTGAGAACTCCT 5296

Qy 5101 aaggaaagtctgaagtcocctgaagacocctggccggtctcctcagctcttccagacacca 5160
Dy 5297 AAGGAAAGTCTGAAGTCCCTGAAGACCTGGCCGGCTTCATCGAGCTCTCCAGACACCA 5356
Qy 5161 agtccactaaggaaatcaatgactaataatgaataaactccaaagtatcctacagagctcca 5220
Dy 5357 AGTCACACTAAGGAATCAATGACTAATGAAAAAAGTACCAAGTATCCTACAGAGCTTCA 5416
Qy 5221 cagccagacctagtgagacaccccaacagctccaagccacagcccaagagaagctcagg 5280
Dy 5417 CAGCCAGACCTTAGTGAGACCCCAACAAAGCTCCAAGCCACAGCCCAAGAGAAGTCTCAGG 5476
Qy 5281 aaagcagacactgaagaagaatttttagcattttaggaacaacacgcctcagcgacaaa 5340
Dy 5477 AAAGCAGACACTGAAGAAGAAATTTTAGCATTTAGGAACAAACGCCATCAGCAGCCAAA 5536
Qy 5341 gccatgcacacaccccaacccagcagtagtgaaagaagaagacatacaacacgtttttggga 5400
Dy 5537 GCCATGTCACACACCCAAACCCAGCAGTAGTGAAGAGAAGACATCAACACCGTTTTTGGGA 5596
Qy 5401 actcagtgcaaaactgagccagccaggaataatttacctggcgaataatagacgactcaaa 5460
Dy 5597 ACTCAGTGCAAGAACTGGACCGCCAGCGAAATTTACCTGGCAGCAATAGACGGCTACAA 5656
Qy 5461 actcgttaaggaaagccagcgtctagaagaactgactggtctcagagagcttttccag 5520
Dy 5657 ACTCGTAAGGAAAGGCCCGCCAGCTCTAGAGAAGTACTAGTGGCTTCAGAGAGCTTTCCAG 5716
Qy 5521 acacatgcaatataccccacagcgtgtagagaaaactccaaaaataactctgcgaaa 5580
Dy 5717 ACACCATGCATGATAACCCACACAGCTGATGAGAAAATACCAAAAAAATACTCTGCAAAA 5776
Qy 5581 tctcgcgaatcagaccagcggaacaccccaacaaacaaagcaacgcgccaagagaagc 5640
Dy 5777 TCTCGGCANTCAGACCCAGCGGACACCCCAACAAACAAAGCAACAGGCCCAAGAGAAGC 5836
Qy 5641 ctcaagaagcagcgtagaggaagaatttttagcattcaggaactcaacccatcagca 5700
Dy 5837 CTCAGAAAGCAGACGCTAGAGGAAGAATTTTAGCATTCAGGAAACTAACACCATCAGCA 5896
Qy 5701 ggcgaagccatgcacacgccttaagcagcagcagtagtggaagagaagacatacaacattt 5760
Dy 5897 GGCAAAAGCCATGCAACGCCCTAAAGCAGCAGTAGGTGAAGAGAAGACATCAACACATTT 5956
Qy 5761 gtgggactccagtggagaactggacctgctaggaattttacctggcagcaagagacgg 5820
Dy 5957 GTGGGACTCCAGTGGAGAACTGGACCTGCTAGGAAATTTACCTGGCAGCAAGAGACGG 6016
Qy 5821 ccaaaactcctaaagaaagccaaagcctctagaagatctggctggcttcaagagctc 5880
Dy 6017 CCACAACTCTAAAGAAAGGCCCAAGGCTCTAGAGAATCTGGCTGGCTTCAAGAGCTC 6076
Qy 5881 ttccagacacacaggttcacactgaggaaatcaatgaccgtagacaaaaatcagaagatccc 5940
Dy 6077 TTCCAGACACCCAGGTTCACACTGAGGAATCAATGACCGGATGACAAAATCACGAAGTATCC 6136
Qy 5941 tgcgaatctccacacacagaccagtcraaaaccccaacaaagctccaagcaacgactcaag 6000
Dy 6137 TGCAAAATCTCCAAACACCAAGCCAGTCAAAACCCCAACAAAGCTCCAAGCAGCAGCTCAAG 6196
Qy 6001 atatacttgggaaagttagtgtaaaagaaggttcctaccagctggcgaagctcacacag 6060
Dy 6197 ATATCTTTGGGAAAGTAGTGTGAAGAAGAGGTCCTACCAGTCGGCAAGCTCACACAG 6256
Qy 6061 acgtcaggaagaccacacacagacacagagagacagcagagatggaagagacatcaaa 6120
Dy 6257 ACGTCAGGGAAGACACACAGACACACAGAGAGACAGCAGGAGATGGAAGAAGACATCAAA 6316
Qy 6121 gcgtttaaggaaatctgcaaaagcaatgctgacccacgcaaaactatggaactgggatgag 6180
Dy 6317 GCGTTTAAAGGAATCTGCAAAAGCAGATGCTGGACCCAGCAAACTATGGAACCTGGGATGGAG 6376

Qy 6181 aggtggccaaagacacacctaaggaagggcccccactactagaagacctgcccggcttcaaa 6240
Dy 6377 AGGTGCCAAGAAACACCTTAAGGAAGAGGCCCAATCACTAGAAGACCTGGCCGGCTTCAAA 6436
Qy 6241 gagctcttcagagacaccagaccacactgaggaaatcaacaactgatgacaaaaactaccaaa 6300
Dy 6437 GAGCTCTTTCCAGACACCAGACCACACTGAGGAATCAACAACTGATGACAAAACTACCAAA 6496
Qy 6301 atagctctcaaatctccaccaccagaatacaatggacactcccaacaagcacaagagggcgg 6360
Dy 6497 ATAGCTCTCAAAATCTCCACCACCAAGTCAATGGACACTTCCAAACAGCAAGAGGCCGG 6556
Qy 6361 cccaaaaacacctttgggaaaggagataatgtggaagagctctcagccctgaagcagctc 6420
Dy 6557 CCCAAACACCTTTGGGGAAGAGGATATAGTGAAGAGCTCTCAGCCCTGAAGCAGCTC 6616
Qy 6421 acacagacacacacacagacaaaatcaccagagatgaggataaaaggacataaacgtgttc 6480
Dy 6617 ACACAGACCCACACACAGACAAAGTACCAGGAGATGAGGATAAAGGCATCAACGTGTTT 6676
Qy 6481 agggaaactgcaaaacagaaaactgacccagcagcaagtgttaactggttagcaagagcgag 6540
Dy 6677 AGGGAACATGCAAAACAGAACTGGACCCAGCAGCAAGTGTAACTGGTAGCAAGAGCGAG 6736
Qy 6541 ccaagaactcctaaaggaaaaagcccaaccccttagaagacttggctggttgaagagctc 6600
Dy 6737 CCAAGAACTCTTAAGGGAAGAGCCCAACCCCTAGAAGACTTGGCTGGCTTGAAGAGCTC 6796
Qy 6601 ttccagacacacgtatgcactgacaaagccacagactcacgagaaaaactccaaaaatagcc 6660
Dy 6797 TTCCAGACACCAAGTATGCACTGACAAAGCCACGACTCAGGAGAAAATACCAAAAATAGCC 6856
Qy 6661 tgcagatctcacacacagaccagctgggtaccccaacaatctcaagccacagctccag 6720
Dy 6857 TGCAGATCTCCACACACAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAGTCCAAG 6916
Qy 6721 agaactctcagaaagcagacgtagaggaagaatccttagcactcaggaaaaagcaacacca 6780
Dy 6917 AGAAGTCTCAGGAAAGCAGACGCTAGAGGAAGAATCCTTAGCATCAGGAAACGAACACCA 6976
Qy 6781 tcagtaggaaagctatggacacaccccaacccagcagaggtgtagagaaaagacatgaaa 6840
Dy 6977 TCAGTAGGGAAGAGCTATGGACACACCAACCAAGCAGGAGGTGATGAGAAAAGACATGAAA 7036
Qy 6841 gcatttatggaaactccagtgacgaataatggacctgcccaggaatttacctggcagcaaa 6900
Dy 7037 GCATTTATGGGAACCTCCAGTGCAGAAATTTGACCTGCCAGGAATTTACTTGGCAGCAAA 7096
Qy 6901 agatggccacaaactcctaaaggaagggcccgagctctagaagacctgctggcttcaaa 6960
Dy 7097 AGATGGCCACAAACTCCTTAAGGAAAGGCCCAAGGCTCTAGAAGACCTGGCTTCAAA 7156
Qy 6961 gagctcttcagagacaccaggcaactgacaagcccaagcagactgtagagaaaaactccaaaaata 7020
Dy 7157 GAGCTCTTCCAGACACCAGGGCACTGACAAAGCCCAAGCAGCTGATGAGAAAACTACCAAAATA 7216
Qy 7021 gctctcaaatctcccaacacagaccagctggacccccagccagcaagcacaagcaagggccc 7080
Dy 7217 GCCTCAAAATCTCCAACACCAACCAAGCCAGTGGACACCCCAAGCAAGCAAGCAAGCCGCC 7276
Qy 7081 aagagaaactcaggaagcagacgtagaggagaattttttagcactcagggaaaaagca 7140
Dy 7277 AAGAGAAACCTCAGGAAAGCAGACGCTAGAGGAAGAATTTTAGCATCAGGAAACGAACA 7336
Qy 7141 ccatcagaggaacaaagccaatggaccccccaaacagcagtagtaagtagagaaaaataatc 7200
Dy 7337 CCATCAGCAGGCAAGCCATGGACACCCCAAAACCCAGCAAGCAAGCAAGCAAGCCGCC 7396
Qy 7201 aacacatttggaaactccagtcagaaactggacctgctaggaataattacctggcagc 7260
Dy 7397 AACACATTTGTGGAAACTCCAGTGCAGAACTGGACCTGCTAGGAATTTTACTGGCAGC 7456
Qy 7261 aagagacagccacagactcctctaaggaagggctgaggctcctagagacacctggttggtcttc 7320

Db 7457 AAGAGACGCCACAGACTCTCTAGGAAAAGCTGAGGCTCTAGAGGACCTGGTTGGCTTC 7516
QY 7321 aaagaactcttcagacacaccaggctcacactgaggaaatcaatgactgatgacaaaatcaaca 7380
Db 7517 AAGAACTCTTCAGACACCCAGGTCACACTGAGGAATCAATGACTGATGACAAATCACA 7576
QY 7381 gaagtatccctgtaaatctccacagcagagctcattccaaaacctcaagaagctccaagcaa 7440
Db 7577 GAAGTATCTCTGTTAAATCTCCACAGCCAGAGTCAATTCAAAACCTCAAGAAGCTCCAAGCAA 7636
QY 7441 agctcaagatccccctggtagaagtggacatgaagaagaagagcccttagcagtcagcaag 7500
Db 7637 AGGCTCAAGATACCCCTGGTGAAGTGGACATGAAGAAGAGGCCCTTAGCAGTCAGCAAG 7696
QY 7501 ctcacacggacatcaggggagactacgcaaacacacacagagcgaacacagagagatagtaag 7560
Db 7697 CTCACAGGACATCAGGGAGACTACGCAACACACACAGAGCCACACAGAGATAGTAAG 7756
QY 7561 agcatcaaaagctttaaggagctccaaaagcagatccttgaccacagcagcaagtgtaaact 7620
Db 7757 AGCATCAAAAGCGTTTAAGGAGTCTCCAAAGCAGATCTCTGGACCCAGCAGCAAGTGAAC 7816
QY 7621 ggtagcagagcagctgaactctgaagaaaagcccgctgctctagaagacctgggt 7680
Db 7817 GGTAGCAGGAGGCGAGCTGGAACCTCGTAAGAAAAGGCCCGTGTCTTAGAAGACCTGGTT 7876
QY 7681 gacttcaaaagagctctctcagcaccaggctcacactgaagagtcaatgactattgacaaa 7740
Db 7877 GACTTCAAAAGAGCTCTCTCAGCACCAGGTCACACTGAAGAGTCAATGACTATTGACAAA 7936
QY 7741 aacacaaaattccctgcaaatctccccaccagaactaacagacactgccacagacaca 7800
Db 7937 AACACAAAATTCCTGCAAAATCTCCCCACCAGAACTTAACAGACACTGCCACGAGCACA 7996
QY 7801 aagagatgcccagacagctcccagaaaagaagtaaaagaggagctctcagcagttgag 7860
Db 7997 AAGAGATGCCCCAAGACACGCTCCAGGAAAGAGTAAAAGAGGAGCTCTCAGCAGTTGAG 8056
QY 7861 aggtctcacgcaaacatcagggcaagcacacacacacacaaagaacacagcagctgtat 7920
Db 8057 AGGCTCACGCAACATCAGGGCAAGACACACACACACACAAAGAACAGCAGCGGTGAT 8116
QY 7921 gagggcatcaaatgtgaagaactgcaagaagaaccccaacccagtagaaggagaa 7980
Db 8117 GAGGGCATCAAAATATTGAAGCAACCTGCAAGAAAGAAACCAACCCAGTAGAAGAGGAA 8176
QY 7981 cccagcagagaagccaaagacacctaaggaagcccaacccctgaaagacctggcc 8040
Db 8177 CCCAGCAGNAGAGCCNAGAGCCCTAAGGAAAAGGCCCAACCCCTGGAAGACCTGGCC 8236
QY 8041 ggcttcacagagctctctgaaacatcaggtcacactcaggaatcactgactgctggcaaa 8100
Db 8237 GGCTTCACAGAGCTCTCTGAAACATCAGGTTCACACTCAGGAATCACTGACTGCTGCCAAA 8296
QY 8101 gcaactaaaataacctggaatctcccacactagaagtggtagaacaccacagcaagacaca 8160
Db 8297 GCCACTAAAATACCCCTGCGAATCTCCCCACTAGAAGTGTGAAGAGAGCCCTCAGCAGTC 8356
QY 8161 aagagcatctcaggacacgtgtgcagaaggtacaagtataaagaagagccttcagcagtc 8220
Db 8357 AAGAGCATCTCAGGACACGTGTGCAGAAGGTACAAGTAAAAGAAGAGCCCTCAGCAGTC 8416
QY 8221 aagttcacacaaacatcaggggaaacccaggtgcagacaaagaacccagcaggtgaagt 8280
Db 8417 AAGTTCAACAAACATCAGGGGAAACCACGGATGCAGACAAAGAACACAGGTGAAGAT 8476
QY 8281 aaaggcatcaagcattgaggaatctgcaaaacagacaccggctccagcagcaagtgtaca 8340
Db 8477 AAAGGCATCAAAAGCATTTGAAGGAATCTGCAAAACACACACCCGGCTCCAGCAGCAAGTGA 8536
QY 8341 actggcagcagagacgcccgaagacaccagggaaagtgcacaaagcctatagaagacctta 8400

Db 8537 ACTGGCAGCAGGAGAGCGGCCCAAGAGCACCCAGGGAAAGTGCCTCCCAAGCCATAGAAACCTA 8596
QY 8401 gctggcttcaaaagacccagcagcaggtccacactgaagaaatcaatgactgatgacaaaacc 8460
Db 8597 GCTGGCTTCAAAAGACCCAGCAGCAGGTTCACACTGAAGAATCAATGACTGATGACAAAACC 8656
QY 8461 actaaataacctgcaaatcatcaccagaactagaagacacccgcaacaagctcaaaagaga 8520
Db 8657 ACTAAAATACCTTGCAAAATCATCCAGAACTAGAAGACACCCGCAACAGCTCAAGAGA 8716
QY 8521 cggcccagagacacgtgcccaagaagttagaagtgaaggaggagctgttagcagttggcaag 8580
Db 8717 CGGCCCAGGACACGTGCCCAGAAAGTAGAAGTGAAGGAGGAGCTGTTAGCAGTTGGCAAG 8776
QY 8581 ctcacacaaacctcaggggagacacgacacacgacaaagacggttaggtgagggcaca 8640
Db 8777 CTCACAAACCTCAGGGGAGACCACGCACACACGCAAGAGCCGTAGGTGAGGCGCAA 8836
QY 8641 ggcaagaagcatttaagcaacctgcaagcggaacgctggacgcagaagatgttaattggc 8700
Db 8837 GGCACGAAAAGCATTTAAGCAACCTGCAAAAGCGGAAGCTGGACGCAGAAGATGTAATTGGC 8896
QY 8701 agcagagacagccaaagacacctaaggaagaaagggcccaacccctggaaagcctggccagc 8760
Db 8897 AGCAGGAGACAGCCAAAGACCCTAAGGAAAAGGCCCAACCCCTGGAAGACCTTGGCCAGC 8956
QY 8761 ttcaagagctcttcaaacacaccagggccacactgaggaactggcaaatggtgctgtgat 8820
Db 8957 TTCCAAAGAGCTCTCTCAACACCAGGCCACACTGAGGAACCTGCGCAATGCTGCTGTGAT 9016
QY 8821 agctttacaagcgtccaaagcaaacacacctgacagtggaacaccttaaaaaatccaga 8880
Db 9017 AGCTTTACAAGCGCTCCAAGCAAAACACCTGCAGCTGGAAGTGGAAAACCTCTTAAAAATATCCAGA 9076
QY 8881 agagttcttcgggcccctaaagtagaacccgtgggagaagtggttaagcaccagagaccct 8940
Db 9077 AGAGTTCTTTCGGGCCCTTAAAGTAGAACCCGTGGGAGACGTGTGTAAGACACAGAGACCT 9136
QY 8941 gtaaatcaaaagcaaaagcaaacacttcctgtccccactgccttcaagaggggaggt 9000
Db 9137 GTAAAATCACAAGCAAAAAGCAACACTTCCTTCGCCCCACTGCCCCTTCAAGAGGGGAGGT 9196
QY 9001 ggcaagaatggaagcgtcacgggaaccaaagagcgtcgctgcatgcccagccagaggaa 9060
Db 9197 GGCAAAAGATGGAAGCGTCAACGGGAACCAAGAGCTGCGGTGTCATGCGCAGCACCAGAGGAA 9256
QY 9061 attgtgagagagctccagccagcaagagagaggttgcctcccagggcagaagcaca 9120
Db 9257 ATTGTGAGGAGCTGCCGCCAGCAGAGAGGAGGTTGCTCCCAGGGCAAGAGGCAAA 9316
QY 9121 tcatccgaacccgtgggtcatcatgaagagaagtttgaggacttctgcaaaaagaattgaa 9180
Db 9317 TCATCCGAACCCGTGTGTCATCATGAAGAGAGAAGTTTCAGGACCTTCGCAAAAAGAAATTGAA 9376
QY 9181 cctgaggaagagctgaacggcaacgcacatgaaaaaccaaaagagaggaacacaaattacaa 9240
Db 9377 CCTGCGGAAGAGTGAACAGCAGCAACGCATGAAAACCAACAAAGAGGAAACAAATTTACA 9436
QY 9241 gactcggctccctgaaataaagggaatatccctgcgctccagacgccaagataagactgag 9300
Db 9437 GACTCGGTCCCTGAAAATTAAGGGAATATCCCTGCGCTCCAGCAGCCCAAGATAAGACTGAG 9496
QY 9301 ggaagacagcaataactgaggtctttgtattagcagaagaatagaataacacagaaat 9360
Db 9497 GCAGAACACGCAAAATACTCAGGCTCTTTGTATTAGCAGAGAAGAATAAGAAATAACACAAAT 9556
QY 9361 gaaagaagcccatgaagacctccccagagatggacattcagaatccagatgagagcc 9420
Db 9557 GAAAAGAACCCATGAAGACCTCCCCAGAGATGGACATTTCAGAATCCAGATGATGAGCC 9616
QY 9421 cggaaacccatcccttagagacaaaagtcactgagaacaaaggtgctgaggtcgtctaga 9480
Db 9617 CGGAAAACCCATACCTTAGACACAAAGTCACTGAGAACAAAGGTCCTTGAGGTCGTCTAGTA 9676

Qy	9481	cagaatgagagctccagcctaagtggtgagagggagcggaggcagaagaagtcgcaag	9540
Db	9677	CAGATGAGAGCTCCAGCGCTTAAGTGGCAGAGGAGCGGAGGCGAGAAGAGTGCGAAG	9736
Qy	9541	gttctcatgcagaaatcagaagggaagagagagagagaaatcagactccatgtgcctg	9600
Db	9737	GTTCTCATGCAGAAATCAGAAGGGAAGAGAGAGAGAAATTCAGACTCCATGTGCCTG	9796
Qy	9601	agatcaagaagaacaaaagccagcctgcagcaagcacttggagagcaaaatcctgtgcag	9660
Db	9797	AGATCAAGAAACAAAAGCCAGCGCTGCAGCAACACTTTGGAGAGCAAAATCTGTGCAG	9856
Qy	9661	agagtaacgcggagtgtaagagggtgcagaaaaatccaaaagagcgtgagacaaatg	9720
Db	9857	AGAGTAACGCGAGGTGTAAGAGGTTGCAGAAAAATCAAGAAGGCTGAGGACAAATGTG	9916
Qy	9721	tgtgtcaagaataaacacagaaagtcagatgagggcagtggaagata	9766
Db	9917	TGTGTCAAGAAATAAACACCAAGAGTCATAGGGACAGTGAAGATA	9962
RESULT 4			
HSWK167A	LOCUS	HSWK167A	11435 bp mRNA linear PRI 31-JAN-1994
DEFINITION	H.sapiens mki67a mRNA (short type) for antigen of monoclonal antibody Ki-67.		
ACCESSION	X65551		
VERSION	X65551.1 GI:415820		
KEYWORDS	antigen; monoclonal antibody.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 11435)		
TITLE	Gerdes Fors,J.		
JOURNAL	Direct Submission		
REMARK	Submitted (11-APR-1992) J. Gerdes, Forschungsinstitut Borstel, Div. Molecular Immunology, Parkallee 22, 2061 Borstel, FRG		
AUTHORS	sequence revised by author 14-JUL-93 and 08-OCT-93		
TITLE	2 (bases 1 to 11435)		
JOURNAL	Schluter,C., Duchrow,M., Wohlenberg,C., Becker,M.H., Key,G., Flad,H.D. and Gerdes,J.		
MEDLINE	The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins		
FEATURES	J. Cell Biol. 123 (3), 513-522 (1993)		
source	94043435 Location/Qualifiers		
exon	1..11435		
exon	/organism="Homo sapiens"		
gene	/db_xref="taxon:9606"		
CDS	/chromosome="10q25-qter"		
	/cell_line="IM9"		
	/clone_lib="lambda gtl1"		
	1..107		
	/number=1		
	108..288		
	/number=2		
	197..8887		
	/gene="mki-67"		
	197..8887		
	/gene="mki-67"		
	/codon_start=1		
	/product="antigen of the monoclonal antibody Ki-67"		
	/protein_id="CAA46520.1"		
	/db_xref="GI:415821"		
	/db_xref="SWISS-PROT:P46013"		
	/translation="MWPTRLVITIKRSGVDGPHRPLSLTCLFGRGIRCDIRIOLPVV SKOCHKTEIHQEAIIHNFSTNPTVNGSVIDEPVLRKHGDVITIDRSFYENESL QNGKSTFPFRKIREQEPARKVRKSRSSDSEGIPLKRRRVSGFGLRPLFDEN LPPNTPLRKGAPTRKSLVMHTPPVLKKIIKEQPPSGKQESGEIHVEYKAQSLVI SPPAPSPRKTPVASDQRRRSCKTAPASSKSKSQTEVPKRGGERVATCLQKRVYSRSQH		

DILQMICSKRRSGSEANLIVAKSWADYVVKLGAKOTQTKVIKHGQPSRMNKRORPAT
PKVPVGHVSQSTGHANSPTIIIGKANTEKRVHPARPVRYVNLNFIINQKMDKDFEVL
SGIAEMFKTPVKEOPOLSTCHIIISNSENLLGKQFOGCTDSEGEPLLPSTSSFGNVE
FSAQNAAKPQSDCKSASPLRQOCIRENGNVAKTPTNTYKMTSLTKTSDTETESKT
PVTYNSRSTEFNRNIQKLPVESKSEETNVEICILARGQKALQORRGEMKEIE
RPFTYRIENIELKENDERKAMKRWGQKCAPMSDITDLKSLDTELMKMDTARGON
LLQOTDHAAPKTKSEKGTIKMPCQSLQPEPINTPTHTKQOLKASLGKVGVBELIAG
KFTRTSGTTHREPADGKSIIRTFKESPKQILDPAARVTGMMKWPPTPKKEASLE
DLAGFKELFQTPGSESMTEKTKIACKSPPPSVDTPTSTKWPKSRKLRKADVEE
BFLAKLITPSAGKAMLTPKPAGGDEKIKAFMGTPVOKLIDLGLTSPKSRLOLPKE
KAQLEDLAGFKELFQTPGHTHEELVAAGTKTKIPCDSPQSDPVDPTSTKQPKRSIR
KADVEGELLACNLMPSSAGKAMHTPKPSVGEKQIIIFVTPVQKLDLENLTSGSKRR
POTPKEEAALEDLTGFKELFQTPGHTHEEVAAGTKTKMPCESSPEADPTSTFRRQ
PKTLEKRDVQKELSALKKLTQSTGTHTDKPGGEDKSIINAFRETKAQKLDPAASV
TKSKRHPKTKKAOPLDLAGKELFQTPVCTDKPTHEKTKIACRQSPDPTPTS
SKPOKSRSLRVQDVEEFAALRKRTPSAGKAMHTPKPAVSGEKNIYAFMGTPVOKLIL
TENLTGSKRRLOTPEKAAQLEDLAGFKELFQTPGHTHEESMTNDKTAACKVSSQDPL
DKNPASSARRLKTSILGKVGVEELIAGKLTQSTGTHHTHTHTHTHTHTHTHTHTHTHT
KOILDSASLTGSKRRQRTPKGSEVPEDLAGFELFQTPSHTKESMNEKTKVSYR
ASQPLVDPTSSKPPQKRSRKADTEEFALFRQTPSAGKAMHTPKPAVGEEDKIN
TFLGTPVOKLDQGNLPGSNRRLOTPEKAAQLEDLTGFKELFQTPCTDNPADTKTT
KKILCKSPQSDPADPTNTKORPKSLKADVEEFLAFRLITPSAGKAMHTPKPAAGV
BEKDINTFVGTPEKLDLGLNLPCKRRPOTPEKAKALEDLAGFKELFQTPGHTHEES
MTDOKITVSCSKSPQDPVKTPSTSSKRLKISLGKVGVEEVLPGKLTQTSKTKTQT
HRETAGDGKSIKAFKESAKQMLDPANYGTGMWRPRTPEKEAQLEDLAGFKELFQTP
DHTBESTDDTKTKIACKSPPPESMDTPTSTRRRPKTPGLKRDIVIELSALKQLTQT
HTDKVPGEDKGINVRETAKOKLDPAASVTGSKRPRTPKGAQPLEDLAGKELFQ
TPVCTDKPTHEKTKIACRSPQDPVGTPTIFKPSKRSRLKADVEESLALRRKTP
SVGRAMDTPKPAGGDEKMAFMGTPVOKLDLPGNLPGSKRWPQTPPEKAAQLEDLAG
FKELFQTPGTDKPTTDEKTKIACKSPQDPVDPTASTQPKRNLKADVEEFLAL
RKRTPSAGKAMDTPKPAVSDKMNITFETPVOKLDLGLNLPCKRRPOTPEKAEAL
EDLVGFKELFQTPGHTHEESMTDDKITEVSCSKSPQSEPKTSRSSKQRLKILPVKVDK
EPLALEDLVDFKELFSAPGHTESMTIDKNTKIPKSPPELTATSTKRCPTRP
EKARALEDLVDFKELFSAPGHTESMTIDKNTKIPKSPPELTATSTKRCPTRP
RKEYKEELSAVERLTQTSQSTHTHKEPASGDEGKIVLKQRAKKPNPVEEPSRRRP
RAPKEAQPLEDTELTSETSGHTQESLTAGKATKIPCESPPVLEVVDVTASTKRLH
RTRVQKVQVEEPSAVKFTQTSGETTDADKAPAGEDKGIKALKESAQTPAPAAAVTG
SRPRAPRESAQAIEDLAGFKDPAAGHTESMTDDTKTKIPKSSPELTATSTSKR
RPRTAQKVEKVEELIAGKLTQTSGETTHTDKPEVGGKGTAKFQPAKRVNDAEDV
IGSRQPRAPKEAQPLEDLASFOELQTSOTPGHTEELANGAADSTFSAKQTPDSCGPL
KISRVLRAKVPVGVVDVTRDPVKSQSNSTSLPPLPFKRGGKSDSVTGCKPLRC
MPAPEEIVVEELPASKQIRVAPRARGKSEVPVIMKRSURTSAKRIEPAEELNSMDK
NKEEHLQDSVPENKQISLRSSRODTEAQOITFVLAERIEINRNKPKMKTSP
MDIQNDPQCARPKIPRDKVTENKRLSARQNESSQPKVAEESGQKSAKVLQKQKG
KGEAGNSDMSCLRSRKTKSQPAASTLESKSVQRTVRSVKRCAENPKKAEDNVCKKIT
TSHRSEDII"

289..367 /gene="mki-67"
/number=3
368..483 /gene="mki-67"
/number=4
484..550 /gene="mki-67"
/number=5
561..596 /gene="mki-67"
/number=6
597..772 /gene="mki-67"
/note="exon 7 is partially excluded by splicing"
/number=8
773..1085 /gene="mki-67"
/number=9
1086..1204 /gene="mki-67"
/number=10
1205..1385 /gene="mki-67"
/number=11
1386..1532 /gene="mki-67"
/number=12

exon	/number=12									
	1533..8377									
	/gene="mki-67"									
	/number=13									
exon	8378..8821									
	/gene="mki-67"									
exon	/number=14									
	8822..11413									
polyA_signal	/number=15									
	8914..8919									
polyA_signal	9449..9454									
	10157..10162									
polyA_signal	11388..11393									
	11414									
BASE COUNT 3802 a 2828 c 2684 g 2121 t										
ORIGIN										
Query Match 84.8%; Score 8286; DB 9; Length 11435;										
Best Local Similarity 100.0%; Pred. No. 0;										
Matches 8286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
Qy	1481	atgagagtgggaatacctttgaaagaagcggtgtctcttggcgccactaagac	1540							
Db	597	ATGAGAGTGGGAATACCTTTGAAAAGAAGCGTGTCTTTGGTGGGCACCTAAGAC	656							
Qy	1541	ctgaactattgatgaataacttgcctcctaatacgcctctcaaaaggggagaagcccaa	1600							
Db	657	CTGAACATTTGATGAAACCTTGCTCTCTAATACGCCCTCTCAAAAGGGGAGAGCCCCAA	716							
Qy	1601	ccaaagaaaagtctctgttaatgcacactccacctgtccttgagaataatcatcaaggaa	1660							
Db	717	CCAAAAGAAAGTCTCTGTAATGCACACTCCACCTGTCTCTGTAAGAAATCATCAAGGAAC	776							
Qy	1661	agcctcaaccatcaggaaacaagagtcagggttcagaataccatgtggaagtgaaggcac	1720							
Db	777	AGCCTCAACCATCAGGAAACAGAGTCAGGTTCAAGAAATCCATGTGGAAGTGAAGGCAC	836							
Qy	1721	aaagcttggttataagccctccagctcctagtctcctaggaaactccagttgccagtgatc	1780							
Db	837	AAAGCTTGGTTATAAGCCCTCCAGCTCCTAGTCTAGGAAACCTCCAGTTGCCAGTGATC	896							
Qy	1781	aagcgcgtagtgcctgcaaaaagccctcgtctccagcagcaaatctcagacagaggttc	1840							
Db	897	AACGCGTAGGTCTGCAAAACAGCCCTGCTTCCAGCAGCAAAATCTCAGACAGAGGTTTC	956							
Qy	1841	ctaagagagggagaaagagatggcaacctgccttcaaaagagagtgtctatcagccgaa	1900							
Db	957	CTAAGAGAGGAGGAGAAAGAGTGGCAACCTGCCCTTCAAAAGAGAGTGTCTATCAGCCGAA	1016							
Qy	1901	gtcaacatgatattttacagatgatattgttccaaaagaagaagtgtgtctcggaagcaa	1960							
Db	1017	GTCAACATGATATTTACAGATGATATGTTCCAAAGAGAAAGTGTGTCTTCGGNAGCAA	1076							
Qy	1961	atctgattgtgcaaatcatatggcgagatgtagtataaaacttggtgcaaaaacaacacaaa	2020							
Db	1077	ATCTGATTTGTCAAAATCATGGCGAGATGTAGTAAAACTTTGGTGCAAAACAAACACAAA	1136							
Qy	2021	ctaaagtcataaacaatggctcctcaaggtcgaatgaacaaaaggcaaaagagacctgcta	2080							
Db	1137	CTAAAGTCATAAACATGGTCTCTCAAGGTCAATGAACAAAGGCAAAAGAACCTTGCTA	1196							
Qy	2081	ctccaaagaagcctgtggcgaaagttcacagttcaatttagtacaggccacgcgaactctc	2140							
Db	1197	CTCCAAAGAAGCCTGTGGCGGAAGTTTACAGTCAATTTAGTACAGGCCACGCCTCTCTC	1256							
Qy	2141	cttggtaccataataatagggaaagctcatactgaaaaagtagtatgtgcctgtcgcacctc	2200							
Db	1257	CTTGTACCATTAATAAGGGAAGAGTCATACTGAAAAAGTACATGTGCTGCTCGACCTC	1316							
Qy	2201	acagagtgtccaacaacttcatctccaacccaaaataatgactttaagggaagatctttcag	2260							
Db	1317	ACAGAGTGCTCAACAACCTTCATTTCCACCACCAAAATGCACTTTAAGGAAGATCTTTCAG	1376							

QY	2261	gaatagctgaaatgtttcaagacccccagtgaaagagcaaccgcagttgacaaagcacatgtc	2320							
Db	1377	GAATAGCTGAAATGTTTCAAGACCCCACTGTAAGGAGCAACCCGCACTTCAACAGCACATGTC	1436							
QY	2321	acatcgctatttcaaatccagagaatttgcttggaataacagtttcaagggaactgatttcag	2380							
Db	1437	ACATCGCTATTTTCAAAATTCAGAGAAATTTGCTTGAAAACAGTTTCAAGGAACATGATTCAG	1496							
QY	2381	gagaagaacctctgctcccccacactcagagagttttggaggaataatgttcttccagtcac	2440							
Db	1497	GAGAAGAACCTCTGTCTCCCACTCAGAGAGTTTGGAGGAATGTGTTCTTTCAGTGCAC	1556							
QY	2441	agaaatcagcaaaacagcgcatctctgataataatgcctctcgaagccctcccttaagacgagc	2500							
Db	1557	AGAAATGCAGCAAAACAGCCATCTGATAAATGCTCTCTGCAAGCCCTCCCTTAAGACGGCAGT	1616							
QY	2501	gtattagagaaaaatggaacgtagcaaaaaacgcccagggaacactactacaaaatgactctc	2560							
Db	1617	GTATTAGAGAAAATGGAAACGTAGCAAAAACGCCAGGAACACCTACAAAATGACTTCTC	1676							
QY	2561	tgagacaaaaacttcagatactagacagagccttcaaaaacagtatccacitgtaaaaca	2620							
Db	1677	TGGAGACAAAACCTTCAGATACTTGACACAGAGCCTTCAAAAACAGTATCCACTGTAAACA	1736							
QY	2621	ggtcaggaaggtctacagagtttcaggaatatatacagaagctcacctgtggaaagttaaagtg	2680							
Db	1737	GGTCAGGAAGGTCTACAGAGTTCAGGAATATACAGAAGCTACCTGTGTGGAAGTAAAGAGTG	1796							
QY	2681	aagaacaaatcacagaaaattgttgatgcatctctaaaaagaggttcagaaggaacactac	2740							
Db	1797	AAGAAAACAAATACAGAAATTTGTTGAGTGCATCTCTAAAGAGGTCAGAAAGCAACATAC	1856							
QY	2741	tacaacaaggagagagagagagatgaaggaataaagaagaccttttgacacataaag	2800							
Db	1857	TACACAAAGGAGAGAGGAGAGATGAAGGAATAGAAAGACCTTTTGGACACATATAAGG	1916							
QY	2801	aaaattgaattaaagaacacgatgaagaatgaagaatgaagaatgaagaatgaagaactt	2860							
Db	1917	AAAAATTTGAAATTAAGAAAACGATGAAAGATGAAGCAATGAAGAGATCAAGAACTT	1976							
QY	2861	gggggagaaaaatgtcaccaatgtctgacctgacagacctcaagagagcttgcctgatacag	2920							
Db	1977	GGGGGAGAAAATGTGACCAATGTCTGACCTGACAGACCTTCAAGAGCTTTCCTGTGATACAG	2036							
QY	2921	aactcatgaagacacgagcgacgtggcgagaatctctccaaacccaagacataagcaag	2980							
Db	2037	AACTCATGAAGACACGGCACGTGGCCAGAAATCTCTCCAAACCCCAAGATCATGCCAAGG	2096							
QY	2981	caccaagagtgaagaaggcaaaatcactaaaaatgccccgtccagtcattacaaccagAAC	3040							
Db	2097	CACCAAGAGTGAGAAAGGCAAAATCACTAAAATGCCCTGCCAGTCATTAACAACCAAGAAC	2156							
QY	3041	caataaaaccccccaac	3100							
Db	2157	CAATAAACACCCCAAC	2216							
QY	3101	tgaagaagaagctccctagcagtcggaagttcacacgagcgtcagggagacacacacaca	3160							
Db	2217	TGAACAAAGAGTCTCTTAGCAGTCTGGCAAGTTTACACGGAGCTCAGGGGAGACCCAGCACA	2276							
QY	3161	cgcacagagagccagcagagagatggaagacatcagaacgttttaaggaggtcttccaaagc	3220							
Db	2277	CGCACAGAGAGCCAGCAGGAGATGGCAAGAGCATCAGAAGCTTTAAGGAGTCTTCCAAAGC	2336							
QY	3221	agatctcgaccacagcagcccggttaactggaatgaagaagtggccaagaacccctaagg	3280							
Db	2337	AGATCTGGACCCAGCAGCCCGTGTAACTTGAATGAAGAGAGTGGCCAAAGACCCCTAAGG	2396							
QY	3281	aagagggccagtcactagaagacgtggcttcaaaagagctcttccagacacccaggtc	3340							
Db	2397	AAGAGGCCCAGTCACCTAGAAGACCTGGCTGGCTTCAAGAGGCTTTCACAGACACAGGTC	2456							

QY 3341 cctctgaggaaatcaatgactgatgagaaactaccaaaatagcctgcaaatctccaccac 3400
Db 2457 CCTCTGAGGAATCAATGACTGATGAGAAATACCAAAATAGCCTGCAAAATCTCCACCAC 2516
QY 3401 cagaatcaatgagactccaacagcacaaaagcaaatgacctgaagagaatctcaggaaaag 3460
Db 2517 CAGAATCAGTGAGCACTCCAAACAGCACAAAGCAATGGCCTAAGGAAGCTTCAGGAAG 2576
QY 3461 cagatgtagagaaatattcttagcactcaggaaactaacacacatcagcagggaagcca 3520
Db 2577 CAGATGTAGAGAGAAATCTTAGCACCTCAGGAATACACACATCAGCAGGGAAGCCA 2636
QY 3521 tgcttacgccccaaaccagcaggagtgatgagaaagacattaaagcattatggaactc 3580
Db 2637 TGCTTACGCCCAACACAGCAGGAGGTGATGAGAAAGACATTAAGACATTTATGGGAATC 2696
QY 3581 caatgcagaactggaactggcaggaaactttacctggcagcaaaagacagctacagactc 3640
Db 2697 CAGTCAGAAACTGGACCTGGCAGGAACCTTACCTGGCAGCAAAAGACAGCTACAGACTC 2756
QY 3641 ctaaggaaaagcccgaggtctctagaagacctggctggtttaaagagctcttccagactc 3700
Db 2757 CTAAGGAAGAAGCCACGAGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACTC 2816
QY 3701 ctggttcacaccgaggaattagtggtctggtgtaaacaccactaaaatcacctcgactctc 3760
Db 2817 CTGGTCACACCGAGGAATTTAGTGGCTGTGTTAAACACCACTAAATATCCCTGCGACTCTC 2876
QY 3761 caagtgcagaccagtggacaccccaacaaagcacaaaagcaagcaccgaagaagtatca 3820
Db 2877 CACAGTCACAGCCAGTGGACACCCCAACAAAGCACAAAGCAAGCACCCCAAGAAAGTATCA 2936
QY 3821 ggaagcagatgtagaggagaactcttagcgtgagggaatctaatgacctcagcaggca 3880
Db 2937 GGAAGCAGATGTAGAGGGAGAACTTTAGCCGTGAGGAACTCTAATGCCATCAGCAGGCA 2996
QY 3881 aagccatgcacgcctaaaccatcagtaggtggaagagaagacatcatatttggg 3940
Db 2997 AAGCCATGCACACGGCTAAACCATCAGTAGGTGAAGAGAAGACATCATATTTGTGG 3056
QY 3941 gaactccagtgcagaaactggacctgacagagaacttaaccggcagcaagagacggccac 4000
Db 3057 GAACCTCAGTCAGAAACTGGACCTTGACAGAGAACTTAACCGGCACAGAGACGGCCAC 3116
QY 4001 aaactcctaaggaagggccaggctctggaagacctgactggctttaaagagctctcc 4060
Db 3117 AAACCTCTAAGGAAGGCCAGGCTCTGGAGAGACCTGACTGGCTTTAAGAGCTCTCC 3176
QY 4061 agaccctgggtcactgaagcagtggtgctgctggcaaaactaactaaaatgcctgcg 4120
Db 3177 AGACCCCTGGTCATACTGAAGAAGCAGTGGCTGCTGGCAAAACTACTAAAAATGCCCTGG 3236
QY 4121 aatcttctccacagaatacagcagacaccccacaaagcaagaagggcagcccaagacac 4180
Db 3237 AATCTTCTCCACAGAATCAGCAGACACCCCAACAAAGCAAGAAGGGCAGCCCAAGACAC 3296
QY 4181 ctttgagaaaaggagctgcagaaaggagctctcagccctgaagagctccacacagacat 4240
Db 3297 CTTTGGAGAAAAGGACGTACAGAGGAGCTCTCAGCCCTGAAGAGCTCACACAGACAT 3356
QY 4241 caggggaaccacacacagataaagtaccaggaggtgagataaaaagcatcaacgcgt 4300
Db 3357 CAGGGGAACCCACACACAGATAAAGTACCAGGAGGTGAGGATAAAGCATCAACGCCGT 3416
QY 4301 ttagggaactgcaaaacagaacttgaccacagcagcaagttaactggtagcaagagcg 4360
Db 3417 TTAGGGAAACTGCAAAACAGAACTTGACCCAGCAGCAAGTGTAACTGGTAGCAAGAGCG 3476
QY 4361 acccaaaactaaggaaggcccaacccttagaagacctggtggtggaagagctct 4420
Db 3477 ACCCAAAAACCTAAGGAAAAGGCCAACCCCTAGAGAGACCTGGCTGGCTGGAAGAGCTCT 3536
QY 4421 tccagacaccagtatgctgacaaagcccccagcactcagcagaaaaactaccaaaatagcct 4480

Db 3537 TCAGACACCAAGTATGACTGTACAAGCCACGACTCACGAGAAAACCTACCAAAATAGCCT 3596
QY 4481 gcagatcaacaacagaccagtcgacacaccaacaagctccaagccacagctccaagagaa 4540
Db 3597 GCAGATCAACAACAGACCCAGTGGACACACCAACAAGCTCCAAGCCACACAGTCCAAGAGAA 3656
QY 4541 gtctcaggaaagtggagctlagaagaagattcttcgcaactcaggaaacgaacaccatcag 4600
Db 3657 GTCTCAGGAAAGTGGAGCTTAGAAGAGAATTTCTTCGCACCTCAGGAACGAACACCATCAG 3716
QY 4601 caggcaaaagccatgcacacaccccaaccagcagtaagtgtgagaaaaacatctacgcct 4660
Db 3717 CAGGCAAAAGCCATGCACACACCCCAACCCAGCAGTAAGTGTGAGAAAAACATCTACGCAT 3776
QY 4661 ttatggaaactccagtgcagaaactgacctgacagaaacttaactggcagcaagac 4720
Db 3777 TTATGGAACTCCAGTGCAGAACTGGACCTGACAGAAACTTTAACTGGCAGCAAGAGAC 3836
QY 4721 ggctacaaactcctaaaggaaaagcccaggtctctagaagacctggctggctttaagagc 4780
Db 3837 GGCTACAAACTCTTAAGGAAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGC 3896
QY 4781 tcttcagacacagaggtcacactgaggaaatcaatgactaaacataaaactgccaaagtag 4840
Db 3897 TCTTCCAGACACGAGGTCACTTGAGGAATCAATGACTTAACGATAAACTGCCAAAGTAG 3956
QY 4841 cctgcataactctcacacaccagacctagacaaaaccagcaagctccaagcagcgctca 4900
Db 3957 CTTGCAAACTCTCACAAACAGACTAGACAAAACCCAGCAAGCTCCAAGCAGGGTCA 4016
QY 4901 agacatccctggggaagtggcgtagaagaagctccctagcagttggcaagctcacac 4960
Db 4017 AGACATCCCTGGGAAAGTGGGCGTGAAAGAGAGCTCTAGCAGTTGGCAAGCTCACAC 4076
QY 4961 agacatcaggagagatcacacacacacacagagcgaacagagagatggttaagagcatga 5020
Db 4077 AGACATCAGGAGAGACTACACACACACACAGAGGCAACAGAGAGATGTTAAGAGCATGA 4136
QY 5021 aagcatttatgagctccaagaagcagacttagactcagcagcaagctctaactggcagca 5080
Db 4137 AGCATTTATGGAGTCTCCAAAGCAGACTTAGACTCAGCAGCAAGTCTTAACCTGGCAGCA 4196
QY 5081 agaggcagctgagaactcctaaaggaaaagtctgaagtcctgaagacctggcgcggttca 5140
Db 4197 AGAGGCAGCTGAGAACTCTTAAGGAAAGTCTGAGTCCCTGAAGACCTGGCCGGCTTCA 4256
QY 5141 ttagagctctccagacaccaagtccacactaaggaaatcaatgactaatgaaaaaactacca 5200
Db 4257 TCAGCTCTTCCAGACACCAAGTCACTAAGGAATCAATGACTAATGAAAAAACTACCA 4316
QY 5201 aagtatcctacagacttcacagccagacctgagacaccccaacaagctccaagccac 5260
Db 4317 AGTATCTTACAGAGTTCACAGCCAGACCTAGTGGACACCCCAACAAAGTCCAAGCCAC 4376
QY 5261 agcccaagagaagctctcaggaagcagacactgaagaagaatttttagcatttaggaac 5320
Db 4377 AGCCCAAGAGAAGTCTCAGGAAGCAGACACTCAAGAAGAATTTTAGCATTTAGGAAAC 4436
QY 5321 aaacgccaatcagcaggcaagccatgcacacaccccaaccagcagtaggtgaagagaaag 5380
Db 4437 AAACGCATCAGCAGGCAAAAGCCATGCACACACCCCAACCCAGCAGTAGGTGAAGAGAAAG 4496
QY 5381 acatcaacacgtttttgggaactccagtcagaaactgacacccagccaggaatttaacctg 5440
Db 4497 ACATCAACACGTTTTTTGGGAACCTCCAGTGCAGAAACTTGGACCCAGGAGAAATTTACCTG 4556
QY 5441 gcagcaatagacggtcacaaactcgttaaggaaaagcccgaggtcttagaagaagactgactg 5500
Db 4557 GCAGCAATAGACGGCTTACAACTCTGTAAGGAAAAGGCCAGGCTCTAGAGAAGACTGACTG 4616
QY 5501 gttcagagagcttttccagacaccatgcactgatacccccacagctgatgagaaaaacta 5560

.Db 4617 GCTTCAGAGAGCTTTTCCAGACACCATTGCACGTGATTAACCCCCACAGCTGATGAGAAAACATA 4676
QY 5561 ccaaaaaatactctgcaaatctccgcaatcagacccagcgagaccccccaacaacacaa 5620
Db 4677 CCAAAAAAATACCTCTGCAAAATCTCCGCAATCAGACCCAGCGGACACCCCAACAACACAA 4736
QY 5621 agcaacggcccaagagaagcctcaagaaaaagacagcgtagaggagaataatttttagcattca 5680
Db 4737 AGCAACGGCCCAAGAGAAGCCTCAAGAAAGCAGACGTAGAGGAAGAAATTTTATGCAATCA 4796
QY 5681 ggaactaaacacatcagcagggcaagccatcacacgcctaaagcagcagtagtggaag 5740
Db 4797 GGAACCTAACACCNATCAGCAGGCAAAAGCCATGCACAGCCCTAAAGCAGCAGTAGAGTGAAG 4856
QY 5741 agaagacatcaacacatttgggggactccagtggaagaaactggagaaactgctaggaatt 5800
Db 4857 AGAAGACATCAACACATTTTGGGGACTCCAGTGGAGAAACTGGACCTCTAGGAAATTT 4916
QY 5801 taactggcagcagagagcgcccaaaactccttaagaaaaaggccaggctctctagaagatc 5860
Db 4917 TACTTGGCAGCAAGAGACGGCCCAAAACTCTTAAGAAAAAGGCCAAGGCTCTAGAAGATC 4976
QY 5861 tggctggcttcaagagctcttcagacacccaggtcacactgaggaatcaatgacagatg 5920
Db 4977 TGGCTGGCTTCAAGAGAGCTCTTCAGHACACCAAGTTCACACTGAGGAATCAATGACCCGATG 5036
QY 5921 acaaaatcacagaagtaacctgcaaatctccacaacccagacccagtcacaaacccccacaa 5980
Db 5037 ACAAAATCACAGAAGTATCTGCAAAATCTCCACAACCCAGACCCAGTCAAAACCCCAACAA 5096
QY 5981 gctccaagcaagactcaagatatccttggggaaagttaggtgtgaaagaagaggtctctac 6040
Db 5097 GCTCCAAAGCAACGACTCAAGATATCTCTTGGGGAAAGTAGGTGTGTAAGAAAGAGGTCCTAC 5156
QY 6041 cagtccggcagctcacacagactcagggaagaccacacacagacacacagagadacagcag 6100
Db 5157 CAGTCGCAAGCTCACACACAGCTCAGGGAAGACCACACAGACACACAGAGACAGACAG 5216
QY 6101 gagatgaaagagcatcaaaagcgtttaaggaattctgaaagcagatgctggacccagcaa 6160
Db 5217 GAGATGAAAGAGCATCAAAAGCGTTTAAAGAAATCTGCAAAAGCAGATGCTGGACCCAGCAA 5276
QY 6161 actatgaaactggatggagaggtggccaaagaacaccttaaggagagggcccaatcactag 6220
Db 5277 ACTATGAACTGGATGGAGGTGGGCCAAGAACACCTTAAGGAAGAGGCCCAATCACTAG 5336
QY 6221 aagacctggccgcttcaagagctcttcagacacacagacacacactgaggaatcaacaa 6280
Db 5337 AAGACCTGGCCGCTTCAAGAGACTCTTCCAGACACCCAGACACACTTGAGGAATCAACAA 5396
QY 6281 ctgatgacaaaactaccaaatagcctgcaaatctccaccaccagaatcaatggacactc 6340
Db 5397 CTGATGACAAAACCTACCAAAATAGCCTGCAAAATCTCCACCACCAGAAATCAATGGACACTC 5456
QY 6341 caacaagcacaagggcgcccaaaaacacttgggggaaaaagggataatagtgaagagc 6400
Db 5457 CAACAAGCAACAAGGCGGCCCAAAACACCTTTGGGGAAAAGGGATATATAGTGGAAAGAGC 5516
QY 6401 tctcagccctgaagcagctcacacagaccacacacacagacaaagtaccagggagatgagg 6460
Db 5517 TCTCAGCCCTGAAGCAGCTCACAGACACACACACAGACAAAGTATACCAGGAGATGAGG 5576
QY 6461 ataaaggcatcaacgtgttcagggaactgcaaaaacagaaactggacccagcgcaagtg 6520
Db 5577 ATAAAGGCATCAACGTGTTCAGGGAAACTGCAAAACAGAAACTGGACCCAGCAGCAAGTG 5636
QY 6521 taactggtagcaagaggcagcaagaactccttaagggaaaaagcccaacccctagaagact 6580
Db 5637 TAACGTGTGCAAGAGGACGCCAAGAACTCTTAAGGGAAAAGCCCAACCCCTAGAAGACT 5696
QY 6581 tggctggcttgaagagctcttcagacacccagtatgctactgacaagccacagactcaag 6640
Db 5697 TGCTGGCTTGAAGAGACTCTTCCAGACACCAAGTATGCTACTGACACAGCCCAACACTCAGC 5756

QY 6641 agaaaactacaaaatagcctgcagatctccacaaccagaccccgtaggggtacccccaaaa 6700
Db 5757 AGAAAACCTACCAAAATAGCCTGCAGATCTCCACAACCCAGACCCAGTGGGTACCCCAACAA 5816
QY 6701 tcttcaagccacagtcacaagagaagtctcaggaaaagcagacgtagagaagaatactcttag 6760
Db 5817 TCTTCAAGCCACAGTCCCAAGAGAAGTCTCAGGAAAGCAGCGTAGAGGAAGAATTCCTTAG 5876
QY 6761 cactcaggaaacgaacaccatcagtagggaaagctaaggacacaccccaaacccagcagag 6820
Db 5877 CACTCAGGAACGACACCATCATAGTAGGAAAGCTATGGACACACCCCAAAACCCAGCAGAG 5936
QY 6821 gtgatgagaagacatgaaagcatttatgggaactccagtgcagaaatttgacactgcacg 6880
Db 5937 GTGATGAGAAGACATGAAGACATTTATGGAACTCCAGTCCAGTAATGGACCTGCCAG 5996
QY 6881 gaaatttacctggcagcaaaaagatggccacaacactcctaaagaaaaagccaggctcttag 6940
Db 5997 GAAATTTACCTGGCAGCAAAAGATGGCCACAAACTCCTTAAGGAAAAGGCCAGGCTCTAG 6056
QY 6941 aagacctggcttcaagagctcttcagacacccagcactgacaagccacacagctg 7000
Db 6057 AAGACCTGGCTGGCTTCAAGAGCTCTTCCAGACACCCAGCAGCTGACAAAGCCACAGACTG 6116
QY 7001 atgagaaaaactacaaaatagcctgcaaatctccacaacccagaccccgtagcacccccag 7060
Db 6117 ATGAGAAAACCTACCAAAATAGCTGCTCAAAATCTCCACAACCCAGACCCAGTGGACACCCAG 6176
QY 7061 caagcacaagcaaacggcccaagagaacactcagggaagcagacgttagagaagaatactt 7120
Db 6177 CAAGCAAAAGCAACGGCCCAAGAGAAACCTCAGGAAAAGCAGACGTAGAGAGAAGAAATTTT 6236
QY 7121 tagcactcaggaaacgaacaccatcagcagggcaaaagcattggacaccccccaaacccagcag 7180
Db 6237 TAGCACTCAGGAACGAAACACCATCAGCAGGCAAAAGCCATGGACACCCCAAAACCCAGCAG 6296
QY 7181 taagtgtatgagaaaaatatcaacacatttggaaactccagtgcaaaactggacaaactggacctgc 7240
Db 6297 TAAAGTATGAGAAAAATATCAACACATTTTGTGGAACATCCAGTGCAGAAAACCTGGACCTGC 6356
QY 7241 taggaatttacctggcagcaagcagacagccacagactcctaaggaaagcctgagctc 7300
Db 6357 TAGGAATTTACCTTGGCAGCAAGAGACAGCCACAGACTCTCTTAAGGAAAAGGCTGAGGCTC 6416
QY 7301 tagaggacctgttgcttcaagaaactctccagacacccaggttcacactgaggaatacaa 7360
Db 6417 TAGAGGACCTGGTTGGCTTCAAGAACTCTTCCAGACACCCAGGTCACACTGAGGAATCAA 6476
QY 7361 tgactgtatgacaaaaatcacagaagtaacctgttaaatctccacagccagctcatctcaaaa 7420
Db 6477 TGACTGATGACAAAATCACAGAAATCTCTGTAAATCTCCACAGCCAGAGTCAATTCAAAA 6536
QY 7421 cctcaagaaactccaagcaaaagctcaagatacccccttggtgaaaagtggacatgaaaagaag 7480
Db 6537 CCTCAAGAGACTCCAAGCAAAAGGCTCAAGATACCCCTTGGTGAAGTGGACATGAAGAAG 6596
QY 7481 agcccttagcagtcagcaagctcacacgggacatacaggggagactcacgcaaacacacacag 7540
Db 6597 AGCCCTTAGCAGTCAAGCAAGCTCACAGGACATCAGGGGAGACTACGGCAAAACACACACAG 6656
QY 7541 agccaacagagatagtagagcatcaaaagcgtttaaggaggtctccaaaagcagatccttgg 7600
Db 6657 AGCCAACAGGAGATAGTAAGAGCATCAAAAGCGTTTAAAGAGTCTTCCAAAAGCAGATCCTCG 6716
QY 7601 acccagcagcaagttaactggtagcagaggcagctgagaactctgtaagaaaaagcccc 7660
Db 6717 ACCCAGCAAGTGTAACTGGTAGCAGGAGGAGCTGGAAGCTCGTAAGAAAAGGCC 6776
QY 7661 gtctcttagaagacctggttgacttcaaaagctctctctcagcaccaggttcacactgaag 7720
Db 6777 GTGCTCTAGAAGACCTGGTTGACTTCAAGAGCTCTTCTCAGCACCCAGGTCACACTCAAG 6836

QY	7721	agtcaatgactattgacaaaacacaaaaattccctctgcaaatctccccccaccagaaactaa	7780
Db	6837	AGTCATGACTATTGACAAAAACAAAAAATTCCTCTGCAATCTCCCCACCAGAACTAA	6896
QY	7781	caaacactccacagacacaaagatgcccaagacacgtcccaggaagaagtaaaag	7840
Db	6897	CAGACTGCCACGACGACAAAGATGCCCAAGACACGTCCCGAAGAAGTAAAG	6956
QY	7841	aggagctctcagcagttgagggctcacgcaaacatcagggcaaaagcacacacacaca	7900
Db	6957	AGGAGCTCTCAGCAGTTGAGGCTCAGCCAAAATCATCGGCAAAAGCACACACACA	7016
QY	7901	agaaaccagcaagcggtgatgaggcatcaaaagtattgaagcaacgtgcaaaagaagaac	7960
Db	7017	AAGAACCACGACGCGTGATGAGGCTCAAAAGTATTGAAGCAACGTGCAAAAGAGAAAC	7076
QY	7961	caaacccagttagaaggaacccacgacgagagaagcccaagacacactaaggaagccccc	8020
Db	7077	CAAACCCAGTAGAAGAGAAACCCAGGAGGAAGGCCAAGACACCTAAGGAAAAGGCC	7136
QY	8021	aacctctggaagacctggccggtctcacagagctctctgaaacatcaggtcacactcagg	8080
Db	7137	AACCCCTGGACAGCTGGCCGGCTTCACAGAGCTCTCTGAAACATCAGTCTACACTCAG	7196
QY	8081	aatactgactctggaagccactaaataacctctgcaattctcccccaactagagtgg	8140
Db	7197	AATCATGACTCTGGCAAAAGCCACTAAATATACCTCTGCAATCTCCCCACTAGAAATGG	7256
QY	8141	tagacacacagacacaaagagcctctcagacacgtgtgcagaaggtacaagtaa	8200
Db	7257	TAGACACCACAGCAGCACAAGAGGCATCTCAGGACACGTGTGCAAGAGGTACAAAGTAA	7316
QY	8201	aagaagagcttcagcagtcgaagttcacacaaacatcaggggaaacccaggtgcagaca	8260
Db	7317	AAGAAGAGCTTCAGCAGTCAAGTTCACACAAATCATCAGGGGNAACCCAGGATCGAGACA	7376
QY	8261	agaaaccagcaggtgaagataaaggcatcaaaagcattgaaggaattctgcaaacacagcac	8320
Db	7377	AAGAACCAGCAGGTGAAGATAAAGGCATCAAAAGCATTTGAAGGAATCTGCAAAACAGACAC	7436
QY	8321	cggtccagcagcaagttgaactgcagcagcagcagcgcgaagacacccagggaaagtg	8380
Db	7437	CGGCTCCAGCACAAGTGTAACTGGCAGCAGGAGAGCGCCCAAGACCCAGGGAAGTGG	7496
QY	8381	cccaagccataagaacacctagcttgcttcaagaccacagcagcaggtcacactgaagaat	8440
Db	7497	CCCAAGCCATAGAAGACCTAGCTGGCTTCAAGACCCACGACGAGTCTACACTGAAGAT	7556
QY	8441	caatgactgatgacaaaaccactaaataacctgcaaatcaataccagaaactagaagaca	8500
Db	7557	CAATGACTGATGACAAAACCACTAAATATACCTTGCAAAATCATCACCAAACTAGAAAGACA	7616
QY	8501	ccgcaacagctcaagagacgcccagacacgtgcccagaaagttagaagtgaaggagg	8560
Db	7617	CCGCAACAGCTCAAGAGACCGCCAGACACGTGCCCCAGAAAGTAGAAGTGAAGGAGG	7676
QY	8561	agctgttagcagttggcgaagctcacacaaacctcagggagacccagcacaccgacaaag	8620
Db	7677	AGCTGTATGACACTTGGCAAGCTCACAAAOCTCAGGGGAGACCGCACACCCACAAAG	7736
QY	8621	agccggttagtgaggcaagcagaaagcatttaagcaacctgcaaaagcggaacgtgg	8680
Db	7737	AGCCGGTAGTGGGGCAAAAGGCACAAAGCATTTAAGCAACGTGCAAAAGCGCAACGTGG	7796
QY	8681	acgcaagaatgataattgacagcagagacacccagacacactaaggaagaaagcccaac	8740
Db	7797	ACGCAAGAATGTAAATTGGCAGCAGGAGACGCCAAGACACCTAAGGAAAAGGCCCAAC	7856
QY	8741	ccctggaagacctggccagcttccaaagagctctctcaaacacacagggccacactgaggaac	8800
Db	7857	CCCTGGAAGACCTGGCCAGCTTCCAGAGCTCTCTCAAAACACAGGCCACACTGAGGAAC	7916
QY	8801	tggcaaatggtgctgtagctttacaagcgtctccaaagcaaacacacctgacagtgga	8860

Db	7917	TGCAAAATGGTGTGTGTAGCTTTACAGCGCTCTCAAAAGCAAAACACCTTGACAGTGGAA	7976
QY	8861	aaccttaaaaaatatcagaagagttcttcggccctaaagttagaacccgttgagacg	8920
Db	7977	AACCTCTAAAATATCCAGAAGAGTTCTTCGGGCCCTTAAGGTAGAACCCGTGGGACCG	8036
QY	8921	tgtgaagcaccagagagccctgtaaaaatcacaagcaaaaagcaaacacttccctgccccac	8980
Db	8037	TGTTAAGCACACAGAGACCTGTAAAATCACAAGCAAAAACAACTTCTCCTGCCCCAC	8096
QY	8981	tgcccttcaagaggagggtggcaaaagatggaagagtcacgggaacccaagagcgtcgct	9040
Db	8097	TGCCCTTCAAGAGGGAGGTGGCAAAAGATGGAAGCGCTCACGGGAACCAAGAGCGTGCCT	8156
QY	9041	gcatgcacacaccagaggaattgtgagagagctgccagccagcaagaagacagaggttg	9100
Db	8157	GCATGCCACACCCAGAGGAATTTGTGGAGGAGCTGCCAGCCACGAAGACAGAGGGTTG	8216
QY	9101	ctcccagggcaagaggaataatcatccgaacccgtgggtcatcatgaagagaggtttgagga	9160
Db	8217	CTCCCAGGCAAGAGGCAAAATCATCCGAACCCGTGTGTATCATGAAGAGAAGTTTGAGGA	8276
QY	9161	cttctgcaaaaagaattgaacctgcggaagagctgaacacagcacatgaataccaaca	9220
Db	8277	CTTCTGCAAAAAGAAATTGAACCTTGCGAAGAGAGCTGAACAGACGACATGAAAACCAACA	8336
QY	9221	agagaaacacaaattacaagactcgtccctgaaataaagggaatatccctcgctccca	9280
Db	8337	AAGAGAAACAAATTTACAAGACTCGGTCTCTGAAAATAAGGGAATATCCTCGCTGCCA	8396
QY	9281	gacgcaagataagactgagcgagagacagcaataaactaggtctttgtattagcaaaa	9340
Db	8397	GAGCCCAAGATAAGACTGAGCGAGAACAGCAATAAATACTGAGGTCTTTGTATTAGCAAAA	8456
QY	9341	gaataaataaacacagaaatgaaaaagcccatgaagacctccccagagatggacattc	9400
Db	8457	GAATAGAAATTAACAGAAATGAAAAGAAAGCCCATGAAGACCTCCCCAGAGATGGACATTC	8516
QY	9401	agaatccagatgatgagcccgaacccataacctagacaaagtcactgagaaacaaa	9460
Db	8517	AGAAATCCAGATGATGAGCCCGGAAACCCATACCTAGAGACAAGTCACCTGAGAACAAA	8576
QY	9461	ggtgcttgaggtctgtacacagaatgagagctcccagcctaaggtggcagagagagcg	9520
Db	8577	GGTGCTTGAGGTCTGTGTAGACAGAAATGAGAGCTCCAGCCCTAAGGTGCGCAGAGAGCG	8636
QY	9521	gagggcagaagatgcgaaggttctcatgcagaatcagaagaagggaaggagagcagaa	9580
Db	8637	GAGGGCAGAAGAGTGGCAAGGTTCTCATGCAAGATCAGAAAGGGAAGAGGACAGGAA	8696
QY	9581	attcagactlccatgctcgtgagatcaagaaagacaaaagccagctgcagcagcactt	9640
Db	8697	ATTACAGCTCCATGTGCTGTAGATCAAGAAAGACAAAGAACCCAGCTTCAGCAAGCACTT	8756
QY	9641	tggagcgaataatgtgcagagagtaacgcggaggtgccaagaggtgtgcagaaataccaa	9700
Db	8757	TGGAGAGCAAAATCTGTGCGAGAGATTAACGCGGAGTGTCAAGAGGTGTGCAGAAAATCCAA	8816
QY	9701	agaagcttaggacaatgtgtgtcaagaaaaatacaacacagaagtcatagggacagtg	9760
Db	8817	AGAAGCTCAGGCAATGTGTGTCTCAAGAAAAATAACACAGAAAGTCATAGGACAGTG	8876
QY	9761	aaqata 9766	
Db	8877	AGATA 8882	

RESULT	5
HSAMAK67	
LOCUS	HSAMAK67
DEFINITION	H.sapiens gene for antigen of the monoclonal antibody Ki-67.
ACCESSION	X74107
	29965 bp
	DNA
	linear
	PRI 25-JUN-1997

VERSION	XT4107.1	GI:1370102	
KEYWORDS	antigen; monoclonal antibody; nuclear protein; tandem repeat.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 29965)		
JOURNAL	Gerdes,J.		
REMARK	Direct Submission		
AUTHORS	Submitted (12-JUL-1993) J. Gerdes, Forschungsinstitut Borstel, Div.		
TITLE	Molecular Immunol., Parkallee 22, 23845 Borstel, FRG		
REFERENCE	sequence revised by author 09-OCT-93		
AUTHORS	2 (bases 1 to 29965)		
TITLE	Duchrow,M., Schluter,C., Wohlenberg,C., Flad,H.D. and Gerdes,J.		
JOURNAL	Molecular characterization of the gene locus of the human cell		
MEDLINE	proliferation-associated nuclear protein defined by monoclonal		
FEATURES	antibody Ki-67		
source	Cell Prolif. 29 (1), 1-12 (1996)		
misc_feature	96177792 Location/Qualifiers		
exon	1..29965		
intron	/organism="Homo sapiens"		
exon	/db_xref="taxon:9606"		
intron	/chromosome="10"		
exon	/cell_line="IM9"		
intron	/map="q25-qter"		
exon	1..729		
intron	/note="CpG rich island"		
exon	75..180		
intron	/number=1		
exon	/evidence=experimental		
intron	/product="antigen of the monoclonal antibody Ki-67"		
exon	181..522		
intron	/number=1		
exon	/evidence=experimental		
intron	523..703		
exon	/number=2		
intron	/evidence=experimental		
exon	/product="antigen of the monoclonal antibody Ki-67"		
intron	545..568		
exon	/note="box 1"		
intron	569..592		
exon	/note="box 2"		
intron	704..3100		
exon	/number=2		
intron	/evidence=experimental		
exon	3101..3179		
intron	/number=3		
exon	/evidence=experimental		
intron	/product="antigen of the monoclonal antibody Ki-67"		
exon	3180..3273		
intron	/number=3		
exon	/evidence=experimental		
intron	3274..3389		
exon	/number=4		
intron	/evidence=experimental		
exon	/product="antigen of the monoclonal antibody Ki-67"		
intron	3390..6958		
exon	/number=4		
intron	/evidence=experimental		
exon	6959..7025		
intron	/number=5		
exon	/evidence=experimental		
intron	/product="antigen of the monoclonal antibody Ki-67"		
exon	7026..9741		
intron	/number=5		
exon	/evidence=experimental		
intron	9742..9787		
exon	/number=6		
intron	/evidence=experimental		
exon	/product="antigen of the monoclonal antibody Ki-67"		
intron	9788..10270		
exon	/number=6		


```
repeat_unit      22166. .22585
/note="Ki-67 no 14"
repeat_unit      22586. .22882
/note="Ki-67 no 15"
repeat_unit      22883. .23239
/note="Ki-67 no 16"
intron           23693. .24569
/number=13
/evidence=experimental
24570. .25013
/number=14
/evidence=experimental
25014. .27009
/number=14
/evidence=experimental
27010. .29601
/number=15
/evidence=experimental
/product="antigen of the monoclonal antibody Ki-67"
27102. .27107
27637. .27642
/note="inactive"
polyA_signal     28345. .28350
polyA_signal     29576. .29581
polyA_signal     8995 a 6446 c 6578 g 7946 t
BASE COUNT
ORIGIN
```

```
Query Match      70.1%; Score 6847; DB 9; Length 29965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6847; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2416  gaagaaatgttcttcagtcacagaatgcagcaaacacccactctgataaatgctct 2475
      |||||||
Db  16847  GGAGGAATGTGTTCTTCAGTGCACAGATGCGCAAAACAGCCATCTGATAAATGCTCT 16906

Qy  2476   gaagccctcccttaagcggcagtgattagagaaaaatggaacgtgcaaaaacgccc 2535
      |||||||
Db  16907  GCAAGCCCTCCCTTAAGACGGCAGTGTATTAGAGAAATGGAACGTAGCAAAACGCC 16966

Qy  2536  aggaacactcaaaatgactctctggagacaaaacttcagatcactgagacagacct 2595
      |||||||
Db  16967  AGGAACACTACAAATGACTTCTCTGGAGACAAAACCTTCAGATACTGAGACAGAGCTT 17026

Qy  2596  tcaaaaacagtatccactgtaaacaggtcaggaagttctacagagttcaggaatatcacg 2655
      |||||||
Db  17027  TCAAAAACAGTATCCACTGTAAACAGGTCAGGAAGGCTCTACAGAGTTTCAGGAATATACAG 17086

Qy  2656  aagctacctgtggaagtgaagtaagtaagaaacaaatacagaaattgttgatgcattccta 2715
      |||||||
Db  17087  AAGCTTACCTGTGGAAGTAAAGAGTGAAGAAACAAATACAGAAATTTGTTGAGTGCATCCTA 17146

Qy  2716  aaaaaggttcagaagcgaacactactactacaaaagagagagaagagagatgaagaaata 2775
      |||||||
Db  17147  AAAAGAGGTCGAAGGCAACACTACTACAAACAAAGGAGAGAGAGATGAAGGAAATA 17206

Qy  2776  gaagaccttttgacacataaaggaataattgaattaaaagaaacagatgaagaatg 2835
      |||||||
Db  17207  GAAAGACCTTTTGACACATATPAGGAAATATTGAAATTAAGAAACAGATGAAAAGATG 17266

Qy  2836  aaagcaatgaagagatcagaacttggggcgagaatgtgcaccataatgtctgacctgaca 2895
      |||||||
Db  17267  AAAGCAATGAAGAGATCAAGAACTTTGGGGCGAGAAATGTGCACCAATGTCTGACCTGACA 17326

Qy  2896  gacctcaagagcttgcttgatatacagaactcagaaagacacggcagcggtggccagaatctc 2955
      |||||||
Db  17327  GACCTCAAGAGCTTCCTCGATACAGAACTCATCAAGACACACGGCAGCTGGCCAGAAATCTC 17386

Qy  2956  ctccaaacccaagatcatgccaagcaccaaagagtgagaaagcgaataatcactaaaatg 3015
      |||||||
Db  17387  CTCCAAACCAAGATCAGCCCAAGGCCAACAAAGATGAGAAAGGGCAAAATCACTAAAATG 17446
```

```
Qy  3016  cctgtccagtcattacaaccagaaccaataaaacaccccccaacacacacacacacagttg 3075
      |||||||
Db  17447  CCTGCCAGTCATTACAACAGAACCAATAAACACCCCAACACACACAAACACAGTTG 17506

Qy  3076  aaggcatcctctgggaaagtaggtgaaagaagagctcctagcagtcaggcaagtccaca 3135
      |||||||
Db  17507  AAGGCATCCTCTGGGAAAGTAGGTGTGAAGAAGAGCTCTAGCAGTCGGCAAGTTTACA 17566

Qy  3136  cggacgtcaggggagaccacgcacacgcacagagagccagcagcgagatggcaagagcatc 3195
      |||||||
Db  17567  CGGAGCTCAGGGAGACACACGCACACGCACAGAGCCAGCAGGAGATGCGAAGAGCATC 17626

Qy  3196  agaactgttaaggagtcctccaaagcagatccttgaccacagcgcggtgtaactggaatg 3255
      |||||||
Db  17627  AGAACGTTTAAAGGAGTCTCCAAAGCAGATCCTTGGACCCAGCAGCCCGTGAACCTGAA 17686

Qy  3256  aagaagtgcccaagaacgccttaaggaagagggcccgagtcactagaagacactggtgcttc 3315
      |||||||
Db  17687  AAGAAGTGCCCAAGAACGCCTTAAGGAAGAGGCCCTAGTCTACTAGAAGACCTGGCTGCTTC 17746

Qy  3316  aaagagctcttcagacacacaggtccctctgaggaatcaatgactgatgagaaaaactacc 3375
      |||||||
Db  17747  AAAGAGCTCTTCCAGACACACAGGTCCCTCTGAGGAATCAATGACTGATGAGAAACTACC 17806

Qy  3376  aaatagcctgcataatctccaccaccagaaatcagtggaactcccaacagcacaagaacaa 3435
      |||||||
Db  17807  AAAATAGCCTGCAAAATCTCCACCACCAGAAATCAGTGGACACTCCAACAACAGCACAAAGCAA 17866

Qy  3436  tggcctaagagaagttctcaggaaaagcagatgtagaggaagaatttttaagcactcaggaaa 3495
      |||||||
Db  17867  TGGCCTTAAGAGAAGTCTTCAGGAAGCAGATGTAGAGGAAGAATTCTTACACTACAGAA 17926

Qy  3496  ctaacaccatcagcagggaaagccatgcttcagcccaacccagcagcaggtgatgagaaa 3555
      |||||||
Db  17927  CTAACACCATCAGCAGGGAAAGCCATGCTTACGCCCAAAACCCAGCAGAGGTGATGAGAAA 17986

Qy  3556  gacattaaagcatttatgggaactccagtcgagaaaactggacctggcaggaactttacct 3615
      |||||||
Db  17987  GACATTAAGCATTTATGGGAATCTCCAGTGCAGAAACTGGACCTGCAGGAACTTTACCT 18046

Qy  3616  ggcagcaaaagacagctacagactcctaagaaaaagcccgaggtctagaagacctggct 3675
      |||||||
Db  18047  GGCAGCAAAAGACAGCTACAGACTCCTTAAGGAAAAGGCCAGGCTCTAGAAGACCTGGCT 18106

Qy  3676  ggccttaagagagctcttcagactcctggtcacaccagagaaattagtgctgctgtaaa 3735
      |||||||
Db  18107  GGCCTTAAAGAGCTCTTCCAGACTCCTGCTCACACCGCAGGAATTAGTGGCTGCTGTAAA 18166

Qy  3736  accactaaaataccctgcgactctccacagtcagaccccgagtcgacaccccccaacagaca 3795
      |||||||
Db  18167  ACCACTAAAATACCCTGCGACTCTCCACAGTCAGACCCAGTGGACACCCCAACACAGCACA 18226

Qy  3796  aagcaacgccaagaagaagatcaggaagcagatgtagaggagagaactcttagcgtgc 3855
      |||||||
Db  18227  AAGCAACGCCCAAGAAGAGTATCAGGAAGCAGATGTAGAGGGAGAACTCTTAGCGTGC 18286

Qy  3856  aggaatcctaagtcctcagcagcgaagccatgcaacccctaaacccatcagtagtgaa 3915
      |||||||
Db  18287  AGGAATCTAATGCCATCAGCAGGCAAGCCATGCACACGCCCTAAACCATCATCAGTAGTGAA 18346

Qy  3916  gagaagacatcatcatatttttggaactccagtcagaaaactggacctggacagagac 3975
      |||||||
Db  18347  GAGAAAGACATCATCATATTTTGGGAATCTCCAGTGCAGAAACTGGACCTGCACAGAGAAC 18406

Qy  3976  ttaacggcgaagaagacggccacaaaactcctaaggaagagagcccgaggtctctggaagac 4035
      |||||||
Db  18407  TTAACCGGCAAGAGACGGCCACAACTCCTAAGGAAGAGGCCAGGCTCTTGGGAAGAC 18466

Qy  4036  ctgactggctttaaagagctcttccagacccctggtctactactgaaagaagcagtggtcgt 4095
      |||||||
Db  18467  CTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTGTACTACTGAAAGACGATGGGTGCT 18526

Qy  4096  ggcaaaactactaaaatgacctgcaaatcttctccaccagaatcagcagacacaccccaaca 4155
```

Db 18527 GGCAAAACTACTAAATGGCCCTGGAAATCTTCCACAGAAATCAGCAGACACCCCAACA 18586
QY agcaagaagcagcccaagacacaccttggagaaaaggagcgctacagaaggagctctca 4215
Db 18587 AGCACAGAAGGCCCAAGACACACTTTGGAGAAAAGGAGCGTACAGAAGGAGCTCTCA 18646
QY gccctgaagaagctcacacagacatcagggggaacccacacacagatataagtaccagga 4275
Db 18647 GCCCTGAAGAAGCTTCACACAGACATCAGGGGAAACCCACACACAGATAAAGTACCAGGA 18706
QY ggtagagataaaagcatcaacgcgcttaggggaaactcgaacacagaacactggaccacga 4335
Db 18707 GGTGAGGATAAAAGCATCAACGGCGTTTATGGGAAACTCCAAACACAGAAACTTGGACCCACGA 18766
QY gcaagttaactggttagcaagagcgaccccaaaactaagaaaaggcccaacccctagaa 4395
Db 18767 GCAAGTGAAGTGGTAGCAGAGAGCCACCCAAAACAACTAAGGAAGAGCCCAACCCCTAGAA 18826
QY gacctggctggctggaaagagctcttccagacacccagtatgcactgacaagccacgact 4455
Db 18827 GACCTGGCTGGCTGGAAGAGCTTCTCCAGACACACAGTATGCATGACAAGCCACGACT 18886
QY cagagaaaactaccataaatagcctgcagatcacacacacagaccagctggcacacaca 4515
Db 18887 CACGAGAAAACCTACCAAAATAGCGCTGCAGATCACAAACAGACCCAGTGGACACCAACA 18946
QY agctccaagccacagctccaagagaagctctcagaaaagtgagctgagagaagtaattctc 4575
Db 18947 AGCTCCAAGCCACAGTCCAAGAGAGTCTCAGGNAAGTGGACGTAGAGAAGAAATTCCTC 19006
QY gcactcaggaaaacacacacacacacagcaggcaaaagccaatgcacacaccccaacacagta 4635
Db 19007 GCACCTCAGGAAACGAACACCATCAGCAGGCAAGCCATGCACACACCCAAACCCAGCAGTA 19066
QY agtggtagaaaaacatctacgcgatttatgggaactccagtcgagaaactggacctgaca 4695
Db 19067 AGTGGTGAGAAAACATCTACGGCATTTATGGGAACCTCCAGTGCAGAAACTGGACCTGACA 19126
QY gagaacttaactggcagcaagagacggctacaactcctcagaaaaggcccgagctcta 4755
Db 19127 GAGAACTTAAGTGGCAGCAAGACGGCTACAACTCTTAAGGAAAGGCCAGGCTCTA 19186
QY gaagacctggctggctttaagagctcttccagacacgaggtcacactcagggaaatcaatg 4815
Db 19187 GAAGACCTGGCTGGCTTTAAAGAGCTTCTCCAGACACAGGTCACACTGAGGAATCAATG 19246
QY actaagcataaaactgccaaagtagcctgcaaatcttcacaacagacacttagacaaaac 4875
Db 19247 ACTAACGATAAAACCTGCCAAAGTAGCCTGCANAATCTTCAACCCAGACCTTAGACAAAAC 19306
QY ccagcaagctccaagcgacggctcaagacatccctggggaagtgggctggaagaagag 4935
Db 19307 CCAGCAAGCTCCAAGCGACGGCTCAAGACATCCTGGGGAAGTGGCGGTGAAGAAGAG 19366
QY ctctcagcagctggcgaagctcaacagacatcagggagagactacacacacacacagag 4995
Db 19367 CTCCTAGCAGTGGCAGCTTCACACAGACATCAGGAGAGACTACACACACACACACAGAG 19426
QY ccaacaggagatggttaagagcatgaagcatttatggagctccaaagcagatcttagac 5055
Db 19427 CCAAACGGAGATGGTAGACGATGAAGCATTTATGGAGTCTCCAAAGCAGATCTTAGAC 19486
QY tcagcgaagcttaactggcagcaagagcgagctgagaactccttaagggaagctctgaa 5115
Db 19487 TCAGCAGCAAGTCTAACTGSCACGACAGGCGAGCTGAGRACTCCTTAAGGAAAGTCTGAA 19546
QY gtccctgaagacctgcccggcttcaatcgagctcttccagacacacgaagtcacataaggaa 5175
Db 19547 GTCCTCTGAAGACCTGGCCGGCTTCATCGAGCTTCTCCAGACACCAAGTCAACATAAGGAA 19606
QY tcaatgactaatgaaaaactaccaagtatcctacagagcttcacagcagacacttagtg 5235
.

Db 19607 TCAATGACTAATGAAAAAACTACCAAAAGTATCTCTACAGAGCTTCACAGCCAGACCTAGTG 19666
QY gacaccccaacaagctctcaagccacagcccaagagaagctctcaggaagcagacactgaa 5295
Db 19667 GACACCCCAACAAGCTCCAAGCCACAGCCCAAGAGAAGTCTCAGGNAAGCAGACACTGAA 19726
QY gaagaatttttagcatttagaaaaaaccgcctacagaggcaagccatgcacacacccc 5355
Db 19727 GAAGAAATTTTAGCATTTTAGGAAACAAGCGCATCAGCAGGCAAGCCATGCACACACC 19786
QY aaaccacagtagtgtagagaaagacatcaaacgcttttttgggaaactccaagtccagaaa 5415
Db 19787 AAACCCAGCAGTAGGTGAAGAAAGAAAGACATCAACACGTTTTTGGGAACCTCCAGTGCAGAA 19846
QY ctgacacagccaggaataattacctggcagcaataagagcgctcacaaaactcgtlaaggaagaag 5475
Db 19847 CTGGACCCAGCCAGGAATTTACCTTGGCAGCANTAGAGCGGTACAAACTCTGTAAGGAAAG 19906
QY gccagggctctagaagaactgaactggcttcagagagcttttccagacacactgcactgat 5535
Db 19907 GCCCAGGCTCTAGAAGAACTGACTGGCTTCAGAGAGCTTTTCCAGACACCATGCACTCAT 19966
QY aacccacagctgatgagaaaactaccacaaaaatactctgcaaatctccgcgaactcagac 5595
Db 19967 AACCCACAGCTGATGAGAAAACCTACCAAAAAATATCTCTGCAAACTCCCGCAATCAGAC 20026
QY ccagcgacaccccccaacacacacaaacgagcccaagagaagcctcaagaagaagcagac 5655
Db 20027 CCAGCGACACCCCAACAACACANAGCAACGGCCCAAGAGAACCTTCANAGAAAGCAGAC 20086
QY gttagagaaataatttttagcattcaggaaaactaacacatcagcaggcaagccatgcactgcac 5715
Db 20087 GTAGAGGAAGAAATTTTAGCATTCAGGAAACTTAACCATCAGCAGGCAAGGCCATGCAC 20146
QY acgcctaaagcagcagtagtgtagaagaaagacatcaacacatttctggggactccaagt 5775
Db 20147 ACGCTTAAGCAGCAGTAGGTGAAGAAAGACATCAACACATTTTGGGGACTCCAGTG 20206
QY gagaacctggacctgctaggaaaatttacctggcagcaagagacgcccacaaaactccclaaa 5835
Db 20207 GAGNAACCTGGACCTGCTTAGGAAATTTTACCTGGCAGCAAGAGACGGCCACAACTCCTTAA 20266
QY gaaaagcccaagcctctagaagatctctggcttcaaaagctcttccagacacccaggt 5895
Db 20267 GAAAGGCCAAGGCTCTTAGAAGATCTGGCTGGCTTCAAGAGCTCTTCCAGACACACAGGT 20326
QY cacactgagaaatcaatgacccgatgacaaaatcacagaagatatacctcgaactcccaaa 5955
Db 20327 CACACTGAGGAATCAATGACCCGATGACAAATCACAGAAGTATCTCTGCAAACTCTCCACAA 20386
QY ccagacccagctcaaaaccccccaacgctccaagcaacgactcaagatactccttggggaaa 6015
Db 20387 CCAGACCCAGTCAAAACCCCAACAAAGCTCCAAGCAACGACTCAAGATATCTTTGGGAAA 20446
QY gtagggtgtagaagaaggtccctaccagtcggcaagctcacacagacgctcagggaaagacc 6075
Db 20447 GTAGGTGTAAAGAAGAGGTCCTACCAGTCGGAAGTCAACAGAGCTCAGGGAAGACC 20506
QY acacagacacacagagagacagcagagatggaaaagacatcaaacgctttaaggaatct 6135
Db 20507 ACACAGACACAGAGAGACAGCAGGAGATGGAAAGAGACTCAAGCGTTTAAAGGAATCT 20566
QY gcaagcagatgctggaccacgcaaaactatggaaactggggtggagaggtggcccaagaaca 6195
Db 20567 GCAAAGCAGATGCTGGACCCAGCAAACTATGGAACCTGGATGGAGAGGTGGCCCAAGAAC 20626
QY cctaagaaagagcccaactcactagaagacctggccggcttcaaaagagctcttccagaca 6255
Db 20627 CCTAAGGAAGAGGCCCAATCCTACTAGAGAGCTGGCCGGCTTCAAGAGCTCTTCCAGACA 20686
QY ccagacacacactgagaaatcaacaactgatgacaaaactaccacaaatagcctcgaactat 6315
Db 20687 CCAGACCCACTGAGGANTCAACAACTGTAGACAAAACCTACCAAAATAGCCTGCAAACTCT 20746

Qy	6316	ccaccaccagaatcaatggaactccaaacagcaagggggggcccaaacacaccttg	6375
Db	20747	CCACCACAGAATCAATGGACACTCCAACACAGCAACAAGGAGGGCGCCAAACACACTTTG	20806
Qy	6376	gggaaagggataagtggaagagctcagccctgaagcagctcacacagaccacac	6435
Db	20807	GGGAAAGGGATATAGTGGAAAGACTCTAGCCCTGAAGCAGCTCACACAGACCACAC	20866
Qy	6436	acagacaagtcaccagagatgaggtataaaggcatcaacgtgttcagggaactgcacaa	6495
Db	20867	ACACAAAATACACAGGATGAGATAAAGGCATCAACGTGTCAGGAAACTGCAAAA	20926
Qy	6496	cagaaactggaccacagcagcaagtgttaacttggtagcaaggcgacgcaagaactcctaag	6555
Db	20927	CAGAAATGGACCCAGCAGCAAGTGTAACCTGTAAGCAAGGACGCAAGAACTCCTAAG	20986
Qy	6556	ggaaagcccaacccctagaagacttgctgcttgaagagctcttcagagacacagta	6615
Db	20987	GGAAAGCCCAACCCCTAGAAAGACTTGCTGGCTTGAAGAGCTCTTCCAGACACCACTA	21046
Qy	6616	tgcactgcaagcccaacgactcacagaaaaactaccaaaatagcctcgaatctcccaa	6675
Db	21047	TGCACCTGCAAGCCACGACTCAGCAGAAAACCTACCAAAATAGCCTGCAGATCTCCAAA	21106
Qy	6676	ccagaccagtggtgtaccccaacaattccaagccacagtcctaagagaagtctcagaaa	6735
Db	21107	CCAGACCCAGTGGGTACCCCAACAATCTCAAGCCAGTCCAAGAGAAGTCTCAGGAAA	21166
Qy	6736	gcagactagaaggaagaatccttagcactcaggaacacgaacacacacagtagggaagct	6795
Db	21167	GCAGACCTAGAGGAAGAATCCTTAGCACTCTAGGAAACGAACACCATCAGTAGGGAAGCT	21226
Qy	6796	atgacacacccaaacagcagcaggtgtatgagaaagacatgaagcatitattgggaact	6855
Db	21227	ATGACACACCCAAACAGCAGGAGTGATGAGAAAGACATGAAAGCATTTATGGGAAC	21286
Qy	6856	ccagtcagaaattggacctgccaggaaatttaacctggcagcaaaagatggccacaact	6915
Db	21287	CCAGTGAGAAATGGACCTGCCAGGAAATTTACTGGCAGCAAAAGATGCCACAAACT	21346
Qy	6916	cctaaggaagggccaggctctagaagacactggcttgcttcaaggagctctccagaca	6975
Db	21347	CCTAAGAAAAGGCCCAGGCTCTAGAAGACCTGGCTGGCTTCAAAGAGCTTCTCCAGACA	21406
Qy	6976	ccaggcactgacagccacgactgtatgaaaaactaccaaaatagcctcaaatctcca	7035
Db	21407	CCAGGCACTGCAAGCCACGACTGTATGAAAAACTACCAAAATAGGCTGCAAAATCTCCA	21466
Qy	7036	caaccagaccagttgacaccccgacgacacaaagcaacggcccaagagaaacctcagg	7095
Db	21467	CAACAGACCCAGTGGACCCCAAGCAGCACAAGCAACGCGCCCAAGAGAAACCTCAGG	21526
Qy	7096	aaagcagcagtagagggaagaatttttagcactcaggaaacgaacaccatcagcaggcaaa	7155
Db	21527	AAAGCAGACGCTAGAGGAAGAATTTTATGACTCTAGGAAACGAACACCATCAGCAGGAAA	21586
Qy	7156	gcaatggacaccccaaacacgacagtaagtgtatgaaaaaatatcaacacatttgtgaa	7215
Db	21587	GCCATGGACACCCCAAAACACGACGAGTAGTGATCAGAAAAATATCAACACATTTGTGGAA	21646
Qy	7216	actccagtgcagaactggacctgactggaatttacctggcagcaagagacagccacag	7275
Db	21647	ACTCCAGTGCAGAAACTGGACCTGCTAGGAAATTTACTTGGCAGCAGACAGACGCCACAG	21706
Qy	7276	actctaagggaaggctgagcctctagaggacctggttggtcttcaagaactctccag	7335
Db	21707	ACTCCTAAGGAAAAGGCTGAGGCTCTAGAGACCTGGTGTGGCTTCAAAGAACTCTTCCAG	21766
Qy	7336	acacaggctcacactggaggaatcaatgactgtatgacaaaatcacagaagtatcctgtaaa	7395
Db	21767	ACACGAGTCACTGAGGAATCAATGACTGATGACAAAATCACAGAAGTATCCTGTAAA	21826

Qy	7396	tctccacagccagagtgcatctcaaacctccaagaagctccaagcgaagagctcaagatcccc	7455
Db	21827	TCTCCACAGCCAGAGTCTATTCAAAACCTCAAGAAGCTCCAAGCAAGGCTCAAGATACC	21886
Qy	7456	ctggtgaagctgacatgaagaagagccctagcagtcacgaacagctcacacgacacatca	7515
Db	21887	CTGCTGAAGTGGACATGGAAGAGAGCCCTTAGCAGTTCAGCAAGCTCACACGGACATCA	21946
Qy	7516	ggggagactacgcaaacacacacacagagccacagagatagtaagagacatcaaacgcttt	7575
Db	21947	GGGAGAGTACGCAAAACACACAGAGCCAACAGGAGATAGTAAGAGCATCAAAAGCGTTT	22006
Qy	7576	aagagtgctccaaagcagatcctggaccagcagccaagtgaactggttagcagaggagca	7635
Db	22007	AAGAGTCTCCAAAGCAGATCCTGGACCACAGCAGCAAGTGAACCTGGTAGCAGGAGCAG	22066
Qy	7636	ctgagaactcgaaggaagccgtgctctagaagacctggtgactcaaaagctc	7695
Db	22067	CTGAGAACTCGTAGGAAGAGCCCGTCTCTAGAAAGACCTGGTTGACTTCAAAGAGCTC	22126
Qy	7696	tcttcagcaccaggttcacactgaagagtccaatgactattgcaaaaaacacaaaaattccc	7755
Db	22127	TTCTCAGCACCAAGTTCACACTGAAGAGTCAATGACTATTGCACAAAAACACAAAAATTCCC	22186
Qy	7756	tgcaaatctccccaccagaactaacagacactgccacagcagcaaaaagatgcccccaag	7815
Db	22187	TGCAAACTCTCCCCACCAGAACTAACAGACACTGCCACGAGCACAAAGAGATGCCCAAG	22246
Qy	7816	acagctcccaggaagaagtataaagagagagctctcagcagttgagaggtcacgcacaaca	7875
Db	22247	ACAGCTCCCAGGAAGAAAGTAAAAAGAGAGGCTCTCAGCAGTTGAGAGGCTCACGCCAAACA	22306
Qy	7876	tcaggcaaaagcacacacacacaaaagaccagcaagcgtgtgtaggggcatcaaaagta	7935
Db	22307	TCAGGGCAAAAGCACACACACAAAAAGACCCAGCGGTGATGAGGGGCATCAAAAGTA	22366
Qy	7936	ttgaagcaactgcgaagaagaaacccacccagtagaagaggaaccagcagggagaagg	7995
Db	22367	TTGAAGCAACTGCAAGAAGAAACCAAAACCCAGTAGAAGAGAACCCACCCAGGAGGAAGG	22426
Qy	7996	ccaagagcacctaaggaagaaagcccaacccctggaagacctggccggtttcacagagctc	8055
Db	22427	CCAAGAGCACCTAAGGAAGAGCCCAACCCCTTGAAGACCTGGCCGGCTTCACAGAGCTC	22486
Qy	8056	tctgaacatcaggttcacactcaggaaatcactgactgctggcgaagccactaaatacccc	8115
Db	22487	TCTGAACATCAGGTCACACTCAGGAATCATCTACTGCTGGCAAGCCACTTAAATATCCC	22546
Qy	8116	tgcaaatctccccactagaagtgtagacaccacagcaagcacaagaagcgatctcagg	8175
Db	22547	TGCGAATCTCCCCACTAGAAGTGGTAGACACCACAGCAAGCACAAAGAGGCATCTCAGG	22606
Qy	8176	acagctgctcgaagtgtaacgtaaaaagagaccttcagcagtgcaagttcacacaaca	8235
Db	22607	ACAGCTGTGCAGAGGTACAAGTAAAGAAAGAGCCCTTCAGCAGTCAAGTTCACAGAAACA	22666
Qy	8236	tcagggaacacccagatgcagacaaaagaaacccagcagctgaagataaaggcatcaagca	8295
Db	22667	TCAGGGGAACCCACGGATGCAGACAAAGAACCCAGCAGGTGAAGATTAAGGGCATCAAGCA	22726
Qy	8296	ttgaaggaaactgtcaaaacagacacccgctccagcagcaagtgtaactggcagcaggaga	8355
Db	22727	TTGAAGGAATCTGCAAAACAGACACCCGGCTCCAGCAGCAAGTGAACCTGGCAGCAGGAGA	22786
Qy	8356	cggccaagagcacccagggaagtgtcccaagcca tagaagacctagctggcttcaaaagac	8415
Db	22787	GGGCCAACAGACCCACGGGAAAGTGCCCAAGCCATAGAAGACTAGCTGGCTTCAAGAC	22846
Qy	8416	ccagcagcaggttcacactgaagaaatcaatgactgatgcaaaaaccactaaaaatccctgc	8475
Db	22847	CCAGCAGAGGTCACACTGAAGAAATCAATGACTGATGACAAACCAACCACTTAAATATCCCTGC	22906
Qy	8476	aaatcatcacagaagactagaagacaccccaacaaactcgaagagagcggccagggacagt	8535

[illegible]

RESULT	6
AL390236/c	
LOCUS	123057 bp DNA linear PRI 30-NOV-2001
DEFINITION	Human DNA sequence from clone RP11-380J17 on chromosome 10, complete sequence.
ACCESSION	AL390236
VERSION	AL390236.21 GI:17381338
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 123057)
JOURNAL	Sycamore,N. Direct Submission Submitted (30-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk requests: clonerequest@sanger.ac.uk
On Dec 5, 2001 this sequence version replaced gi:17017713.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw:,
SWISSPROT; Tr, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
RP11-380J17 is from the library RPc1-11.2 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-380J17 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP11-380J17 is at 123057 in this
sequence. The true left end of clone RP11-85C15 is at 103775 in
this sequence. The true right end of clone RP11-4C20 is at 2000 in
this sequence.

FEATURES
SOURCE

```

source
1:123037
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-380J17"
/clone lib="RPC1-11.2"

```

BASE COUNT	ORIGIN
30696 a	29821 c
29797 g	32743 t

Query Match	53.5%	Score 5227;	DB 9:	Length 123057;
Best Local Similarity	99.8%	Pred. NO. 0;		
Matches 5777; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

Qy	3475	gaattcttagcactcaggaaactaacaccatcagcagggaaagccatgcttctgcgccc	3534
Db	123057	GAATTCCTAGCACTCAGGAACTAACACCATCAGCAGGGAAGCCATGCTTACGCCAA	122998

Qy 3535 ccagcaggaggtgatgagaaagacattaaagcattttgggaactccagtgagaaaactg 3594
 |||||
 Db 122997 CCAGCAGGAGGTGATCAGAAAGACATTAAAGCATTTTGGAACTCCAGTCGAGAAACTG 122938
 |||||

Qy	3595	gacctggcaggaaactttacctctggcagcaaaaaagacagctacagactcctaaagaaaagccc	3654
Db	122937	GACCTGCGAGGAACCTTTACCTGGCAGCAAAAGACAGCTACAGACTCTTAAAGAAAAGGCC	122878

Qy	3655	caggcttagaagacctggctggtccttaaaagagctcttccagactccttggtcacaccgag	3714
Dδ	122877	CAGGCTCTAGAAGACCTGGCTGGCTTTAAAGACTCTTCCAGACTCCTGGTCACACCGAG	122818

Qy	3715	gaattagtgctgctggtataaacccactaaaaataacctctgcgactctccaacagtcaagaccaca
Db	122817	GAATTACTGGTGCTGCTGTAAACACCACATAAATATCCCTCGACTCTCCAAGTCAGACGCCA

Qy	3775	gtggacaccccaacagcacacaaagcaacgacccaagaagatctatcaggaaagcagatgta	3834
Db	122757	GTGGACACCCCAACAGCACAAAGCAACGACCCNAGAGAAGTATCAGGAAGACGAGTGTA	122698

Qy 3835 gagggagaactcttagctgaggaatctaatgcatcagcaggcaaaagccatgcacacg 3894
|||||
Db 122697 GAGGAGAACTCTTAGCTGAGGAATCTAATGCCATCAGCAGGCAAGCCATGCACAGC 122638
Qy 3895 cctaaaccatcagtaggtgagagaaagacatcatatttttgggaactccagtcag 3954
|||||
Db 122637 CCTAAACCATCAGTAGGTGAAGAGAAAGACATCATATTATTGTGGGAACCTCCAGTGAG 122578
Qy 3955 aaactggacctgacagagaacttaaccggaagcagcagagagcgcacaaactcctaagaa 4014
|||||
Db 122577 AAACGTGGACCTGACAGAACTTAACCGCGCAAGACGCGCCAAACTTCCTAAGGAA 122518
Qy 4015 gagggccaggtcttgaagacctgactggctttaagagctcttccagacccttggtcat 4074
|||||
Db 122517 GAGGCCAGGCTCTGGAAGACCTGACTGSCCTTAAAGAGCTCTTCCAGACCCCTGGTCAT 122458
Qy 4075 actgaagaagcagtggtgcttggcaaaactactaaatgccttgcaattcttccaca 4134
|||||
Db 122457 ACTGAAGAAGCAGTGGCTGCTGGCAAACTACTAAATGCCCTGGCAATCTTCTCCACCA 122398
Qy 4135 gaatcagcagacaccccaacagcacagaagcagcccaagacacacctttggaagaaag 4194
|||||
Db 122397 GAATCAGCAGACACCCCAACAAAGCAAGAAAGCGCAGCCCAAGACACCTTTTGAGAAAGG 122338
Qy 4195 gacgtacagaagaggtctcagccctgaagagctcacacagacatcaggggaaaccaca 4254
|||||
Db 122337 GACGTACAGAAGGAGCTCTAGCCCTGAAGAAGCTCACACAGACATCAGGGGAACCCACA 122278
Qy 4255 cacacagataaagtaccaggggtgaggaataaaagatacaacgcgttagggaaactgca 4314
|||||
Db 122277 CACACAGATAAAGTACCAAGGAGGTGAGGATAAAGCATCAACGCGCTTTAGGGAACCTGCA 122218
Qy 4315 aaacagaactggacccagcagcagtgtaactggtagcagagagcaccacaaactaag 4374
|||||
Db 122217 AAACAGAAACTGGACCCAGCAGCAAGTGTAACTGGTACAGAGGACCCCAAAACTTAAG 122158
Qy 4375 gaaaaggccaaaccttagagacctggtggtggaagagctcttcagacacacagta 4434
|||||
Db 122157 GAAAAGGCCAACCCCTAGAGACCTGGCTGGCTTGAAGAGCTCTTCCAGACACCACTA 122098
Qy 4435 tgactgcagaagccacagactcacagaaactacccaaatagcctgcagatcacacaaca 4494
|||||
Db 122097 TGCACTGCAAGGCCACACACTCACAGAAACTACCAAAATAGCCTGCAGATCACAACTCA 122038
Qy 4495 gaccagtgacacaccacaaagctcccaagccacagtcacagagagctcaggaagtg 4554
|||||
Db 122037 GACCCAGTGGACACACCAACAAAGCTCCAAGCCACAGTCCAAGAGAGTCTCAGGAAAGTG 121978
Qy 4555 gacgtagaagaagaattcttcgactcaggaaacgaaacaccatcagcaggcaaaagccatg 4614
|||||
Db 121977 GACGTAGAAGAAGAATTCTTGCACTCAGGAACCAACACCATCAGCGCAAAAGCCATG 121918
Qy 4615 cacacacccaaacagcagtgtaagtgtgagaaaaacatacgaatttatgggaactcca 4674
|||||
Db 121917 CACACACCCAAACCCAGCAGTAGTGGTGAGAAAAACATCTACGCAATTTATGGGAACCTCCA 121858
Qy 4675 gtgcagaactggacctgacagagaaacttaactggcagcagagagcgtacaaactcct 4734
|||||
Db 121857 GTGCAGAAACTGGACCTGACAGAGAACTTAAGTGGCAAGAGAGAGCGGTACAAACTCCT 121798
Qy 4735 aaggaaaaggccaggtcttagaagacctggtggtttaaagagctcttcagacacaga 4794
|||||
Db 121797 AAGGAAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACGA 121738
Qy 4795 ggtcacactgaggaataactaactaacgataaaaaactgcgaagtgctgcaaatcttca 4854
|||||
Db 121737 GGTCACTAGGAAATCAATGACTAACGATAAAACTGCAAAAGTAGTACCTGCAAACTCTTCA 121678
Qy 4855 caaccagacctagacaaaacccagcagctcgaagcagcagcgtcaagacatccctggg 4914
|||||
Db 121677 CAACCAGACCCAGACAAAACCCAGCAAGCTCCAAGCGACGGCTCAAGACATCCCTGGGG 121618

Qy 4915 aaagtggcggtgaagaagagctcctagcttggaagctcacacagacatcagagag 4974
|||||
Db 121617 AAAGTGGCGTGAAAGAAAGAGAGCTCTAGCAGTTGCAAGCTCACACAGACATCAGGAGAG 121558
Qy 4975 actacaacacacacacagagccaaacagagagatggttaagagcatgaaagcatttatggag 5034
|||||
Db 121557 ACTACACACACACACAGAGGCCAAAGAGATGTTAAGAGCATGAAAGCATTTATGGAG 121498
Qy 5035 tctccaaagcagatcttagaactcagcagaagctctaaactggcagcaagagcagctgaga 5094
|||||
Db 121497 TCTCCAAAGCAGATCTTAGACTCAGCAGCAAGTCTAAGTGGCAGCAAGAGGAGCTGAGA 121438
Qy 5095 actcctaagggaagctcctgaagctcctgaagacctggcggcgtctcatagagctcttccag 5154
|||||
Db 121437 ACTCTAAAGGAAAGTCTGAAGTCCCTGAAGACCTGGCGGGCTTATCAGAGCTCTCCAG 121378
Qy 5155 acaccaagtccacactaaaggaatcaatgactaatgaaaaaactaccaaagtatctctacaga 5214
|||||
Db 121377 ACACCAAGTCACTAAGGAATCAATGACTAAGCAAAAAAATACCAAAAGTATCCTACAGA 121318
Qy 5215 gcttcacagcagaccttagtgacaccccccaacaaagctccaagccacagcccccaagagaat 5274
|||||
Db 121317 GCTTCACAGCCAGACCTAGTGGACACCCCAACAGCTCCAAGCCACAGCCCAAGAGAGCT 121258
Qy 5275 ctcaggaagcagacactgaagaagaatttttagcatttaggaacaaacgccaatcagca 5334
|||||
Db 121257 CTCAGGAAAGCAGACACTGAAGAAGAAATTTTAGCATTTTAGGAAACAAACGCCATCAGCA 121198
Qy 5335 ggcaaaagccatgcacacacccccaaacagcagtaggtgaagagaaagacataaacagcttt 5394
|||||
Db 121197 GGCAAAAGCCATGCACACACCCCAACAGCAGTAGGTGAAGAGAAAGACATCAACACGTTT 121138
Qy 5395 ttgggaactccagtcagaaactggaccagccagagaaatttacctggcagcaataagacgg 5454
|||||
Db 121137 TTGGGAACCTCCAGTCGAGAAACTGGACCAGCCAGCGGAAATTTACCTGGCAGCAATAGACCG 121078
Qy 5455 ctacaactcgttaaggaaaagcccaggctctagaagaactgactggtctcagagagctt 5514
|||||
Db 121077 CTACAAACTCGTAAGGAAAAAGGCCAGGCTCTAGAAGAACTGACTGGCTTCAGAGAGCTT 121018
Qy 5515 ttccagacacacatgcactgataacccccacagctgatgagaaaaactacccaaaaatactc 5574
|||||
Db 121017 TTCCAGACACCATGCACTGATTAACCCACGACTGATGAGAAAACTACCAAAAAATACTCTC 120958
Qy 5575 tqcaaatctccgcaatcagaccagcgagaccccccaacaaacaaacaaagcaacgccccag 5634
|||||
Db 120957 TGCAANTCTCCCATCAGACCCAGCGGACACCCCAACAAACAAAGCAACGGCCCAAG 120898
Qy 5635 agaagcctcaagaagcagacgttagaggaagaatttttagcattcagggaactaacaacca 5694
|||||
Db 120897 AGAAGCCTCAAGAAAGCAGACAGCTAGAGGAAGAAATTTTAGCATTCAGGAAACTAACACCA 120838
Qy 5695 tcagcaggcaaacatgcacacgcctaaagcagcagtaggtgaagagaaaagacatcaac 5754
|||||
Db 120837 TCAGCAGCCAAAGCCATGCACACGCTTAAAGCAGCAGTAGGTGAAGAGAAAGACATCAAC 120778
Qy 5755 acatttggggactccagtgagaaactggacctgctaggaatttacctggcagcagcaag 5814
|||||
Db 120777 ACATTTGGGGAGCTCCAGCTGGAGAAACTTGACCTGCTAGGAAATTTACTTGGCAGCAAG 120718
Qy 5815 agacggccacaactcctaaagaaaggccaaggtcttagaagatctggtggttcaaa 5874
|||||
Db 120717 AGACGCCACAACCTCTTAAAGAAAAGGCCAAGGCTCTAGAAGATCTGGCTGGCTTCAAA 120658
Qy 5875 gagctcttcagacacccaggtcacactgaggaataatgacccgatagaacaaatcacagaa 5934
|||||
Db 120657 GAGCTCTTCCAGACACCCAGGTCACTGAGGAATCAATGACCGATGACAAAAATCACAGAA 120598
Qy 5935 gtatctcgcaaatctccacaacacagccagctcaaaaccccccaagctcccaagcaacga 5994
|||||
Db 120597 GTATCTGCAAAATCTCACACACAGACCCAGCTCAAAAACCCCAACAGCTTCCAGCAACGA 120538
Qy 5995 ctcaagatatccttggggaaagtaggtgtgaagaagaggttctaccagtcggaagctc 6054

|||||
Db 120537 CTCAAGATATCCCTTGGGGAAGTAGTGTTGAAAGAGAGAGTCTTACCAGTCGGCAAGCTC 120478
QY 6055 acacagacgtcaggggaagacacacagacacacagagagacagagagatgaaagagc 6114
|||||
Db 120477 ACACAGACGTCAGGGAAGACACACAGACACACAGAGAGACACAGAGAGATGGAAGAGC 120418
QY 6115 atcaagcgtttaaggaatctgcaagcagatgctggaccagcaaaactatggaactggg 6174
Db 120417 ATCAAGCGCTTTAAGGAATCTGCAAGACAGATGCTGGACCAGCAAACTATGGAACCTGGG 120358
QY 6175 atggagagtgccagaagaaacacctaaggaagagggcccaatcaatagaaagacctggccggc 6234
|||||
Db 120357 ATGGAGAGTGGCCAGAACACCTTAAGGAAGAGGCCCAATCACTAGAAAGACCTTGGCCGGC 120298
QY 6235 ttcaagagagctcttcagagacacacagaccacactgaggaatcaacaactgatgacaaaact 6294
Db 120297 TTCAAGAGCTCTTCAGACACACAGACCACACTGAGGAATCAACAAGTATGATGACAAAAC 120238
QY 6295 accaaatagctgcgaaatctccaccaccagaatcaatggacactccaacaagcacaagg 6354
Db 120237 ACCAAAATAGCTTGCAAATCTCCACCACCAGAATCAATGGACACTCCAACAAGCACAAGG 120178
QY 6355 agcggcccaaaacacctttggggaagaaaggtatagtggaagagctctcagccctgaag 6414
Db 120177 AGCGGGCCCCAAAACACCTTTGGGGAAAGGGATATAGTGGAAAGAGCTCTCAGCCCTGAAG 120118
QY 6415 cagctcacagacacacacacagacaaagtaccagagagatgagataaaggatcaac 6474
Db 120117 CAGCTCACAGACCCACACACAGACAAAGTACCAGGAGATGAGGATTAAGGCATCAAC 120058
QY 6475 gtgttcagggaactgcaaaacagaactggaccgcagcaagtgttaactggtagcaag 6534
Db 120057 GTGTTCAGGGAAACTGCAAAACAGAACTGGACCACAGCAAGTGTAACTGGTAGCAAG 119998
QY 6535 agcagccaaagaactcctaagggaagaaagcccaaccctagaagacttggctgttgaag 6594
Db 119997 AGCAGCCGAAGAACTCCTTAAGGGAAAGGCCAACCCCTAGAAAGACTTGGCTTGGTAA 119938
QY 6595 gagctcttcagacacacagtatgactgacaaagccacagactcacagagaaaactaccaaa 6654
Db 119937 GAGCTCTCCAGACACCAATATGACTGACAGACCCACGACATCATGAGAAAATACCAAA 119878
QY 6655 atagctgcgaatctccacaacacagaccagctgggttacccccacaatcttcaagccacag 6714
Db 119877 ATAGCTGCAGATCTCCAAACACAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAG 119818
QY 6715 tccaagagaagcttcaggaagcagacgttagaggaagaatccttagcactcaggaaacga 6774
Db 119817 TCCAGAGAAGTCTCAGGAAGCAGACGCTAGAGGAAGAATCCTTAGCCTCAGGAAGCA 119758
QY 6775 acaccatcagtaggggaagctaTggacacaccccaacccagcagaggtgtgatgagaagac 6834
Db 119757 ACACCATCAGTAGGGAAGCTATGGACACACCCCAACCCAGCAGGAGGTGATGAGAAGAC 119698
QY 6835 atgaagcatttatggaaactccagtcgagaaattggacctgccaggaatttacctggc 6894
Db 119697 ATGAAAGCATTTATGGGAATCCAGTGCAGAAATTTGGACTTGCAGGAAATTTACCTGGC 119638
QY 6895 agcaaaagatggccacaactcctaaggaaagggcccaagctctagaagacctggctggc 6954
Db 119637 AGCAAAAGATGGCCCAAACTCCTTAGGAAAGAGGCCACCGCTCTAGAAAGCTTGGCTGGC 119578
QY 6955 ttcaagagctcttcagacacacagcactgacaagcccaacccagcactgatgagaaaactacc 7014
Db 119577 TTCAAAAGAGCTCTTCAGACACACAGCAGCTGACAAACCCACGACTGATGAGAAAATACC 119518
QY 7015 aaaaatagctgcnaatctccacaacacagaccagctggacacccccagcaagcaagcaaa 7074
Db 119517 AAAATAGCTGCAAAATCTCCAAACACAGACCCAGTGGACACCCACAGCAAAAGCAA 119458
QY 7075 cggcccaagaaacctcaggaaagcagacgttagaggaagaatttttagcactcaggaaa 7134
|||||

Db 119457 CGCCCCAAGAGAAAACCTCAGAAAAGCAGAGCTAGAGGAAGAAATTTTATGACACTCAGGAAA 119398
QY 7135 cgaacaccatcagcagcagcaaaagccatggacaccccccaaacccagcagtaagtgtgatgagaaa 7194
|||||
Db 119397 CGAAACCATCAGCAGGCAAAAGCCATGGACACACCAACCAACCCAGCAGCTAAGTGATGAGAAA 119338
QY 7195 aatatacacacatttgtgaaactccagtcgagaaactccagctgcctaggaatttacct 7254
Db 119337 AATATCAACACATTTGTGGAACCTCCAGTGCAGAAACTGGACCTGTGTAGCAAAATTTACCT 119278
QY 7255 ggacgaagaagacagccacagactcctaaggaagaaagctgaggtctctagaggacctgggtt 7314
Db 119277 GGCAGAAAGAGACAGCCACAGACTCCTTAAGGAAAAGGCTGAGGCTCTAGAGGACCTGGTT 119218
QY 7315 ggcctcaaaagaactcttcagagacacaggttcacactgaggaatcaatgactgatgacaaa 7374
Db 119217 GCGTTCAAGAACTCTTCCAGACACCCAGGTACACTGAGGAATCAATGACTGATGACAAA 119158
QY 7375 atcacagaagtaTcctgttaaatctccacagccagagtcattcaaaaacctcaagaagctcc 7434
Db 119157 ATCACAGAAGTATCTGTAAATCTCCACAGCCAGAGTCAATTCAAAAACCTCAAGAAGCTCC 119098
QY 7435 aagcaaaagctcaagataccccctgtggaagtggaacatgaaagaagagcccttagcagtc 7494
Db 119097 AAGCAAAAGCTCAAGATACCCCTGGTGAAGTGGACATGAAGAAGAGCCCTTAGCAGTC 119038
QY 7495 agcaagctcacacgacatcaggggagactacgcaaacacacacagagccaacagagat 7554
Db 119037 AGCAAGCTCACAGGACATCAGGGGAGACTACGCAAAACACACACAGAGCCACAGAGAT 118978
QY 7555 agtaagagcatcaaaagcgtttaaggagctccaaaagcagatcctgtgaccacagcagcaagt 7614
Db 118977 AGTAACAGCATCAAGCGTTTAAAGGAGTCTCCAAGACAGATCCTGGACCAGCAGCAAGT 118918
QY 7615 gtaacttggtagcaggaagcagctgagaactcgttaaggaaaaagggccgctctagaagac 7674
Db 118917 GTAACCTGGTAGCAGGAGGAGCTGAGAATCTGTAAGAAAAGGCCGCTCTCTAGAAAGAC 118858
QY 7675 ctggttgacttcaaaagctcttctcagcagcaggttcacactgaaagttcaatgactatt 7734
Db 118857 CTGTTTGACTTCAAAAGAGCTCTCTCAGCACCCAGGTTCACACTGAAAGTCAATGACTATT 118798
QY 7735 gacaaaaacacaaaatttccctgcnaattccccccaccagaactaacagacactgcacag 7794
Db 118797 GACAAAACACAAAAATTCCTTGCAATCTCCCCACCAGAACTAACAGACACTGCCACG 118738
QY 7795 agcaaaagagatgccccaaagacactccaggaagaagtaaaagagagctctcagca 7854
Db 118737 AGCAAAAGAGATGCCCCAAGACACCTGCCAGGAAAGAGTAAAGAGGAGCTCTCAGCA 118678
QY 7855 gttgagagctcacgcaaaacatacagggcaagcacacacacacacacacacacagcaagc 7914
Db 118677 GTTGAGAGGCTCACGCNAACATCAGGGCAAGGACACACACACACACAAAGACCAAGC 118618
QY 7915 gbtgatggggcatcaaatatttgagcaacgtgcnaagaaagaaacccacccagtagaa 7974
Db 118617 GGTGATGAGGGCATCAAGTATTGAAAGCAACGTCGAAAGAAACCAACCAACCCAGTAGAA 118558
QY 7975 gaggaacccagcaggaagaagccaaagacacctaaggaagaaagcccaacccctggagac 8034
Db 118557 GAGGAACCCAGCAGGAGAAGGCCAAGAGCCCTTAAGGAAAAGGCCCAACCCCTTGGAAAGC 118498
QY 8035 ctggccggcttcacagagctctctgaacatacaggttcacactcaggaatacactgactgct 8094
Db 118497 CTGGCCGGCTTCACAGAGCTCTTGAAACATCAGGTTCACACTCAGGAATCAGTACTGCT 118438
QY 8095 ggcaaaagccactaaaataccctgcnaattccccccactagaagtgttagacaccacagca 8154
Db 118437 GGCAAGGCCACTAAAAATACCTTCCGAAATCTCCCCCACTAGAAAGTGGTAGACACACAGCA 118378
QY 8155 agcaaaagagcactctcaggaacaggtgtgcagaaggtacaagtaaaagaagaccttca 8214
Db 118377 AGCACAAAGAGGCATCTCAGGACACGCTGTGCAGAGGTACAAGTAAAGAAAGAGCCTTCA 118318

QY 8215 gcagtcagtctacacaaacatcaggggaaacccacgagtgatgcagacaaagacacagct 8274
 Db 118317 GCAGTCAAGTTACACAAACATCAGGGGAAACCCACGGATGCAGACAAAAGAACCCAGCT 118258
 QY 8275 gaagataaaggcatcaaaagcattgaagaatctgcaaaacagacacccgctccagcagca 8334
 Db 118257 GAAGATAAAGGCATCAAAAGCATTTGAAGGAATCTGCAAAACAGACACCGGCTCCAGCAGCA 118198
 QY 8335 agtgaactggcagcagcagcgcgcaagagcaccacgaggaagtgcccaagccatagaa 8394
 Db 118197 AGTGTAACTGGCAGCAGAGAGCGGCAAGAGACCCAGGAAAGTGCCCAAGCCATAGAA 118138
 QY 8395 gacctagctggcttcaaaagaccacgacgagtcacactgaagaatcaatgactgatgac 8454
 Db 118137 GACCTAGCTGGCTTCAAAAGACCCAGCAGGTCACACTGAAGAATCAATGACTGATGAC 118078
 QY 8455 aaacccataaataaccctgaaatcatcaccagaaactagaaagacacgcgaacagctca 8514
 Db 118077 AAACCCACTAAATACCTGCAAAATCATCCAGAACTAGAAAGACACCGCAACAAAGCTCA 118018
 QY 8515 aagacacgcccagcagcagctgcccagaaagtgaagtgaagagagctgttagcagtt 8574
 Db 118017 AAGACACGGCCAGACACAGCTGCCAGAAAGTAGAAGTGAAGGAGAGAGCTGTTCACAGTT 117958
 QY 8575 ggcaagctcacacaaacctcagggggagaccgcacacccgacacaaagagccggtaggtgag 8634
 Db 117957 GGCAAGCTCACAAACCTCAGGGGAGACCGGCACACCGCAAAAGAGCGGTAGTGAG 117898
 QY 8635 gcaaaagcagcaaaagcatttaagcaacctgcaaaagcggaaagctgagcagcaaaatgta 8694
 Db 117897 GGCAAAAGCAGCAAAAGCATTTAAGCAACCTGCAAAAGCGGAAGCTGGAGCGCAAGATGTA 117838
 QY 8695 attgcacagagacacccaaagacacctaaggaagaaagcccaacccctgaagacacgtg 8754
 Db 117837 ATTGCACAGAGACGCCAAGAGACACTTAAGGAAAGGCCCAACCCCTGGAAAGATCTG 117778
 QY 8755 gccagctccaagagctctcacaacacagccacactgaggaactggaaatggtgct 8814
 Db 117777 GCCAGCTTCCAAGAGCTCTCTCAAAACACAGGCCACACTGAGGAACCTGGCAATGCTGCT 117718
 QY 8815 gctgatactttacaagcgcctcaaaagcaaacacactgcagctggaacacctctaaaaata 8874
 Db 117717 GCTGATAGCTTTTACAAAGCGCTTCAAAAGCAAAACACTGCACAGTGGAAAAACCTCTAAAAATA 117658
 QY 8875 tcagaagagttcttcgggcccctaaagtgaacccgtggagacgtgtgaagcaccaga 8934
 Db 117657 TCCAAAGAGTCTTCGGGGCCCTAAAGTAGAACCCGTGGGAGAGCTGGTAGACACCCAGA 117598
 QY 8935 gaccctgtataatcacaagaacaaagcaacacttccctgccccactgccccttcaagagg 8994
 Db 117597 GACCCTGTAAAAATCACAAGCAAAAGCAACACTTCCCTGCCCCCACTGCCCTTCAAGAGG 117538
 QY 8995 ggaggtggcaaatggaagcgtcacgggaaccaaagagcgtgcgtgcagccagcacca 9054
 Db 117537 GGAGGTGCAAAAGATGGAAGCGTCAACGGGAACCAAGAGAGCTGCGTGTGATGCCAGCACCA 117478
 QY 9055 gaggaattgtggagagctgcagccagcaagaagcagagaggtgtgtccagggccaaga 9114
 Db 117477 GAGGAATTTGGAGAGACTGCCACCCAGCAGCAAGAGAGGGTTGCTCCCAAGGCCAAGA 117418
 QY 9115 ggcaaatcatccgaaccctgggtcatcatgaagagaagtgtgaggaacttctgcataaaga 9174
 Db 117417 GGCAAAATCATCGAACCCGTGGTGCATCATGAAGAGAAGTTTGAGGACTTCTGCAAAAAAGA 117358
 QY 9175 attgaacctgcgaagagctgaacgcaacacacatgaaaaaccaacaaagagagaaacaaa 9234
 Db 117357 ATTGAACCTGCGGAAGAGCTGAACAGCAACGACATGAAACCAACCAAGAGGAAACACAAA 117298
 QY 9235 ttacaagactcggtccctgaaaaaagg 9262
 Db 117297 TTACAAGACTCGGTCCCTGAAANAATAGG 117270

RESULT 7
 AL355529/c
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-85C15 on chromosome 10, complete sequence.
 ACCESSION AL355529
 VERSION AL355529.21 GI:16972861
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 140466)
 AUTHORS Clark,G.
 TITLE Direct Submission
 JOURNAL Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; ENBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10
 RP11-85C15 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOT: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone RP11-85C15. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true left end of clone RP11-264E18 is at 138467 in this sequence. The true right end of clone RP11-380J17 is at 2000 in this sequence.

FEATURES
 Location/Qualifiers
 1..140466
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-85C15"
 /clone_lib="RPCI-11.1"
 33521..33577
 /note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
 113053
 /note="Tandem repeat. Forced join. Gap size estimated to be approximately 300bp by restriction digest"
 36826 a 30925 c 31089 g 41626 t
 misc_feature
 unsure
 BASE COUNT 36826 a 30925 c 31089 g 41626 t
 ORIGIN

Query Match 29.2%; Score 2855; DB 9; Length 140466;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3055; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db	899	GCACCTCAGAAACGAAACCATCAGCAGCAAGCAATGCACACACCCAAACCCAGCAGTA	840
Qy	4636	agtgtgagaaaaacatctacgcatttatgggaacttcagtcagagaaactgacctgaca	4695
Db	899	AGTGTGAGAAAAACATCTACGCATTTATGGAACTTCAGTCGACAGAACTGGACCTGACA	780
Qy	4696	gagaacttaactggcagcaagagacggtctacaaaactcctaaggaaaagcccgactcta	4755
Db	779	GAGAACTTAACCTGGCAGCAAGAGACGGCTACAAACTCCTAAGGAAAAGCCCGAGCTCTA	720
Qy	4756	gaagacctggctgtttaagagagcttctccagacacacagaggtcacactgaggaatcaatg	4815
Db	719	GAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACAGAGGTTCACACTCAGGAATCAATG	660
Qy	4816	actaacgataaaactgcacaaagttagctgcgaatcttcacaaccagactagacaaaaac	4875
Db	659	ACTAACGATAAACTGCCAAGTAGCCTTGCAAACTTTCACAAACAGACCCAGACAAAAAC	600
Qy	4876	ccagcaagctccaagcggctcaagacatccctcctggggaagtggcggtgaaagaagag	4935
Db	599	CCAGCAAGCTCCAAGCAGCGGCTCAAGACATCCCTGGGGAAAGTGGCGTGAAGAAGAG	540
Qy	4936	ctcctagcagttggcaagtcaacagacatacaggagagactacaacacacacacagag	4995
Db	539	CTCCTAGCAGTTGGCAAGCTCACACAGACATCAGAGAGACTTACACACACACACAGAG	480
Qy	4996	ccaacagagatggttaagagcatgaagcatttatggagctctccaaagcagactctagac	5055
Db	479	CCAACAGAGATGTTAAGACATGAAGACATTTATGGAGTCTCCAAAGCAGATCTTAGAC	420
Qy	5056	tcagcagcaagtctaactggcagcaagcaggcagctgagaactcctaaaggaaagtctgaa	5115
Db	419	TCAGCAGCAAGTCTTAACCTGGCAGCAAGAGCGCAGCTGAGAACTCCTTAAGGAAAGTCTGAA	360
Qy	5116	gtccctgaagacctggcggcttcacagctcttcagacacacaaagtcacactaagaaa	5175
Db	359	CTCCCTGAAGACCTGGCCGGCTTTCATCGAGCTCTTCCAGACACCAAGTCCACACTAAGGAA	300
Qy	5176	tcaatgactaatgaaaaactaccaaagtaacctacagagcttcacagccagacctagtg	5235
Db	299	TCAATGACTTACGAAAAAATACCAAAAGTATCTTACAGAGTTTCACGCCAGCCCTAGTG	240
Qy	5236	gacaccccaagaagctccaagccacagcccaagagaagtctcaggaaaagcagacactgaa	5295
Db	239	GACACCCCAACAAGCTCCAAGCCACAGCCCAAGAGAAGTCTCAGGAAAGCAGACACTGAA	180
Qy	5296	gaagaatttttagcatttaggaacaacacgcctatcagcaggcaagccatgcacacacccc	5355
Db	179	GAAGAAATTTTAGCAATTTAGGAAACAAACCGCCATCAGCAGGCAAAAGCCATGCACACCC	120
Qy	5356	aaaccagcagtaggtgagagaaaagacatcaacacgctttttgggaactccagtcagaaa	5415
Db	119	AAACCAGCAGTAGGTGAGAGAGAAACACATCAACAGCTTTTGGGAACCTCCAGTCGAGAAA	60
Qy	5416	ctggaccagccaggaatttacctggcagcaaatagcggctacacgaactcgttaaggaaa	5474
Db	59	CTGGACCAGCCAGGAAATTTTACCTGGCAGCAATAGACGGCTACAAACTCGTAAAGGAAAA	1

RESULT	8				
LOCUS	AX147668	733 bp	DNA	linear	PAT 08-JUN-2001
DEFINITION	Sequence 1 from Patent WO0136629.				
ACCESSION	AX147668				
VERSION	AX147668.1	GI:14346724			
KEYWORDS	synthetic construct.				
SOURCE	artificial sequence.				
ORGANISM	1 (bases 1 to 733)				
REFERENCE	Gerdes,J., Scholzen,T. and Wohlenberg,C.				
AUTHORS	Transfer compounds, the production and the use thereof				
TITLE					

JOURNAL	Patent: WO 0136629-A 1 25-MAY-2001;				
FEATURES	Forschungszentrum Borstel (DE)				
source	1. 733				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note=" Kon21-DNA-Insert, enthaltend einen Teilbereich der				
	fur das Ki-67 Protein kodierenden Sequenz und				
	Restriktionsstellen"				
BASE COUNT	271 a	144 c	199 g	119 t	
ORIGIN					
Query Match	6.4%; Score 629; DB 6; Length 733;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 629; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	9108	gccaagagcaaatcatccgaaccctggctcatcatgaagagaagtgtgagactctgc	9167		
Db	62	GGCAAGAGCAAAATCATCCGAACCCGTGCTCATATGAAGAGAAGTTTGAGGACTTCTGC	121		
Qy	9168	aaaaaattgaacctgcggaagagctgaacagcaacgacatgaaaaccaaaaagaga	9227		
Db	122	AAAAAGAAATTGAACCTGCGGAAGAGCTGAACAGACACGACATGAACCAACCAAGAGGA	181		
Qy	9228	acacaaattacaagactcgtccctgaaataaagggaatatccctcgcgtccagacgccca	9287		
Db	182	ACACAAATTACAAGACTCGTCCCTCGAAATAAGGGAATATCCCTCGCTCCAGACGCCA	241		
Qy	9288	agataagactcagggaacacagcaataactcaggtctctgtattatgcggaagaataga	9347		
Db	242	AGATAAGACTGAGGCAGAGAACAGCAATAACTGAGGTCTTTGTATTAGCAGAGAATAAG	301		
Qy	9348	aataaacagaaatgaaaagaccccatgaagacctcccccagagatgacattcagaatcc	9407		
Db	302	AATAAACAGAAATGAAAAGAGCCCATGAAGACCTCCCCAGAGATGGACATTTCAGAATCC	361		
Qy	9408	agatgatgagcccggaaccccatacctagagacaaaagtcactgagaaacaaaaggtgctt	9467		
Db	362	AGATGATGAGCCCGGAACCCCATACCTAGAGACAAAGTCACTGAGACAAAGGTTGCTT	421		
Qy	9468	gaggtctgtagaagaaatgagagctcccagcctaaagtggcagagagagcggaggcca	9527		
Db	422	GAGGTCTGTACAGAGAATGAGAGCTCCAGGCTTAAGGTGGCAGAGAGAGCGAGGGCA	481		
Qy	9528	gaagatgcgaaggttctcatgcagaatcagaaaagggaagagagagcaggaattcaga	9587		
Db	482	GAAGAGTGCGAAGGTTCTCATCGAANTCAGAAAGGGAAGGAGAGGAGGAATTCAGA	541		
Qy	9588	ctccatgtcctgagatcaagaaagacaaaagccagcctgcagcagcacctttggagag	9647		
Db	542	CTCCATGTGCTTGAGATCAAGAAAGACAAAAGCCAGCCTGCAGCAAGCACTTTGGAGAG	601		
Qy	9648	caaatctgcgagagataacgcgagtgctcaagaggtctcagaaggttcagaaaaatccaaagaagc	9707		
Db	602	CAAACTGTGTCAGAGAGTAAACGCGGAGTGTCAAGAGGTGTGCAGAAAATCCAAAGAAGGC	661		
Qy	9708	tgagacaatgtgtgtccaagaaaaataa	9736		
Db	662	TGAGGACAATGTGTGTCAAGAAAATAA	690		

RESULT	9				
LOCUS	AX341136	568 bp	DNA	linear	PAT 10-JAN-2002
DEFINITION	Sequence 1383 from Patent WO0196388.				
ACCESSION	AX341136				
VERSION	AX341136.1	GI:18137118			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				

REFERENCE 1 (sites)
AUTHORS Jiang,Y., Harlocker,S.L. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196388-A 1383 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source l..568
BASE COUNT 201 a 144 c 130 g 90 t 3 others
ORIGIN /organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 4.7%; Score 456; DB 6; Length 568;
Best Local Similarity 99.8%; Pred. No. 1.4e-244;
Matches 506; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7243 ggaattaccctggcagcagagacagccacagactcctaaggaaaaagctgaggctcta 7302
Db 1 GGAATTTACCTGGCAGCAAGAGACAGCCACAGACTCCTAAGGAAAAAGGCTGAGGCTCTA 60
QY 7303 gaggacctggtgcttcaaaagaactctccagacaccagggtcacactgagaatcaatg 7362
Db 61 GAGGACTGGTGGCTTCAAGAACHCTTCCAGACACCAAGGTCCACTGAGGAATCAATG 120
QY 7363 actgatgacaaaatacacagaagtatcctgtataatctccacagccagagtcattcaaaacc 7422
Db 121 ACTGATGACAAAATCACAGAAGTATCCTGTATAATCTCCACAGCCAGGGTCATTCAAAACC 180
QY 7423 tcaagaagctccaagcaaggctcaagataccctcctggtgaaagtgcacatgaagaagag 7482
Db 181 TCAAGAAGCTCCAAGCAAAAGGCTCAAGATACCCCTGGTGAAGTGGACATGAAAGAAGAG 240
QY 7483 ccctagcagtcagcagctcacagcagcatcaggggagactacgcaaacacacacagag 7542
Db 241 CCCTAGCAGTCAGCAAGCTCACCGGACATCAGGGGAGACTACGCAAAACACACAGAG 300
QY 7543 ccaacaggagatagtagaagcagcatcaaacgctttaaggagctctccaaagcagatccctggac 7602
Db 301 CCAACAGGAGATAGTAAGAGCATCAAAAGCGTTTAAAGGAGTCTCCAAGCAGATCCTGGAC 360
QY 7603 ccagcagcagtgtaactggtagcagagcagcagcagcagcagcagcagcagcagcagcag 7662
Db 361 CCAGCAGCAAGTGTAACTGGTAGCAGGAGCGAGCTGAGAACTCGTAAAGAAAAGGCCCGT 420
QY 7663 gctctagaagacctggtgacttcaaaagagctctctcagcaccaggtcacactgaagag 7722
Db 421 GCTCTAGAGACCTGGTTGACTTCAAAAGAGCTCTTCTCAGCACCAGGTCCACTGAAGAG 480
QY 7723 tcaatgactattgacaaaaacacaaaa 7749
Db 481 TCAATGACTATTGACAAAAACACAAAA 507

RESULT 10
G33252/c 318 bp DNA linear STS 30-JUN-1997
LOCUS
DEFINITION human STS SHGC-52534, sequence tagged site.
ACCESSION G33252
VERSION G33252.1 GI:2226556
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 318)
Myers,R.M.
Unpublished (1997)
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine

Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu
Primer A: TTTCAGCCAGCCAGCAAGTCTT
Primer B: ACCTTTGGGAAAGGGATA
STS size: 226
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds
Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul
Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3
Prepared with primer pairs derived from W87807 -- Unigene.
FEATURES
Source 1..318
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="10"
STS 92..317
primer_bind 92..111
primer_bind complement(298..317)
BASE COUNT 47 a 69 c 87 g 114 t 1 others
ORIGIN

Query Match 2.5%; Score 245; DB 11; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6368 cacccttgggaaaaaggatatagtgaaagctctcagccctggaagcagctcacacaga 6427
Db 318 CACCTTTGGGAAAAAGGGATATAGTGGAAAGAGCTCTCAGCCCTGGAAGCAGCTCACACAGA 259
QY 6428 ccacacacacagacaaaagTaccaggagatgagataaaaggcatcaacgtgttcaggaaa 6487
Db 258 CCACACACACACACAAAGTACCAGGAGATGAGGATAAAGGCATCAACGTGTTCCAGGAAA 199
QY 6488 ctgcaaaacagaaactggaccagcagcagcaagtgttaactggttagcagagcgccagaa 6547
Db 198 CTGCAAAACAGAAACTGGACCAGCAGCAAGTCTAACTGTAGCAAGAGAGAGCCCAAGAA 139
QY 6548 ctccctaaggaaaaagcccaacccctagaagacttggtgcttgaagagctcttcaga 6607
Db 138 CTCCTAAGGGAAGAAAGCCCAACCCCTAGAAAGACTTGGCTGGCTTGAAGAAGAGCTCTCCAGA 79
QY 6608 cacca 6612
Db 78 CACCA 74

RESULT 11
G03247/c 332 bp DNA linear STS 19-OCT-1995
LOCUS
DEFINITION human STS WI-1011, sequence tagged site.
ACCESSION G03247
VERSION G03247.1 GI:720205

KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE human random genome wide STSs created from sheared whole human DNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 332)
 TITLE Whitehead Institute/MIT Center for Genome Research; Random Genome
 JOURNAL Hudson, T.
 REFERENCE 2 (bases 1 to 332)
 TITLE Whitehead Institute/MIT Center for Genome Research; Physically
 Mapped STS
 JOURNAL Hudson, T.
 COMMENT GDB_DSEG: D10S1249
 Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu
 Primer A: CTTCTGACCTGTTTGCAGT
 Primer B: CTTTCAGTGCACAAATGCAG
 STS size: 201
 PCR Profile:
 Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng
 Primer: each 5 pM
 dNTPs: each 4 mM
 Tag Polymerase: 0.025 units/ul
 Total Vol: 20 ul
 Buffer:
 MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCL: 10 mM
 pH: 9.3.
 Location/Qualifiers
 1..332
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="791_D_7"
 65..265
 65..84
 primer_bind complement(246...265)
 primer_bind 76 a 74 c 67 g 115 t
 BASE COUNT 76 a 74 c 67 g 115 t
 ORIGIN
 STS
 primer_bind
 primer_bind
 BASE COUNT
 ORIGIN
 Query Match 2.0%; Score 199; DB 11; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.4e-99;
 Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2416 ggaggaatgtgttttcagtcgacagaaatgcagcaaacagccatcgtgataaaatgctct 2475
 Db 279 GGAGGAATGTGTTCTTCAGTCAGACAGATGCAGCAAAACAGCCATCTGATAAATGCTCT 220
 QY 2476 gcaagccctcccttaagcgcgagtgattagagaaatggaaacgtagcaaaaacgccc 2535
 Db 219 GCAAGCCCTCCCTTAAGACGGCAGTGATTAGAGAAATGGAAACGTAGCAAAAACGCC 160
 QY 2536 aggaacacctaacaatgacttctctgagacaaaacttcagatactgagacagagcct 2595
 Db 279 GGAGGAATGTGTTCTTCAGTCAGACAGATGCAGCAAAACAGCCATCTGATAAATGCTCT 220

Db 159 AGGAACACCTACAAATGACTTCTCTGGAGACAAAACCTTCAGATACTGAGACAGAGCCT 100
 QY 2596 tcaaaaacagtatccactg 2614
 Db 99 TCAAAAACAGTATCCACTG 81
 RESULT 12
 LOCUS AX336882/c 409 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 7391 from Patent WO0194629.
 ACCESSION AX336882
 VERSION AX336882.1 GI:18127601
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrikan, S., Soppet, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 7391 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 Location/Qualifiers
 source 1..409
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 98 a 90 c 84 g 137 t
 ORIGIN
 Query Match 2.0%; Score 198; DB 6; Length 409;
 Best Local Similarity 100.0%; Pred. No. 5.1e-99;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 493 gacagtaccgcagatgactcaaaagacagtggtgctcaggaacaaactaatgtctatcc 552
 Db 198 GACAGTACCGCAGATGACTCAAAAGACAGTGTGTCAGGGACAACTAATGTTCAATCC 139
 QY 553 tcagaacatcgtggacgtaaatggcagaaatgcagctgacccattctctgggatttaa 612
 Db 138 TCAGAACATGCTGGAGCTAATGCGAAATGCAGCTGATCCCATTTCTGGGATTTAAA 79
 QY 613 gaaattccagcgttaaatagtagccgttatgagaaatgagagctcttcccaactaca 672
 Db 78 GAAATTCAGCGTTAAATAGTAGCCGTTATGGAGAAATGAAGTCTGTCCCACTACA 19
 QY 673 caatgtcttgacaatagc 690
 Db 18 CAATGTCTTGACAAATAGC 1
 RESULT 13
 LOCUS AX072812 158 bp DNA linear PAT 25-JAN-2001
 DEFINITION Sequence 3284 from Patent WO0102568.
 ACCESSION AX072812
 VERSION AX072812.1 GI:12583165
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 158)
 AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
 Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
 Lanson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M.,
 Labat, I., Lesnikowicz, D., Kita, D., Garcia, V. and Strache-Crain, B.
 TITLE Human genes and gene expression products
 JOURNAL Patent: WO 0102568-A 3284 11-JAN-2001;
 CHIRON CORPORATION (US) ; HYSEQ, INC. (US)

```
FEATURES
  source      Location/Qualifiers
             1. .158
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
  misc_feature 1. .158
             /note="n = A,T,C or G"
BASE COUNT   54 a 40 c 34 g 27 t 3 others
ORIGIN

Query Match      1.1%; Score 110; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4699 aacttaactggcagcagacgctctacaaacctcctaagaaagcccgaggtctctagaa 4758
|||||
Db 11 AACTTAACCTGGCAGCAGACGCGCTACAAACCTCTTAAGGAAAGGCCAGGCTCTAGAA 70
|||||

QY 4759 gacctggctggctttaagagctcttcacagacacgaggtcacacctgagga 4808
|||||
Db 71 GACCTGGCTGGCTTTAAAGAGCTCTTCACAGACACGAGGTCACACTGAGGA 120
|||||

RESULT 14
LOCUS      HSK167
DEFINITION H. sapiens DNA for Ki-67 antigen 5'-region (exon 1 & 2).
ACCESSION X94762
VERSION   X94762.1 GI:1944550
KEYWORDS  Ki-67 gene; monoclonal antibody.
SOURCE    human.
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Gerdes,J.
          Direct Submission
          Submitted (08-JAN-1996) J. Gerdes, Molecular Immunology,
          Forschungszentrum Borstel, Parkallee 22, D- 23845, Borstel, FRG
          2 (bases 1 to 14041)
REFERENCE 1
  AUTHORS  Gerdes,J.
  TITLE    Sequence of the human Ki-67 protein gene 5' and promoter region
  JOURNAL  Unpublished
  COMMENT  On Apr 18, 1997 this sequence version replaced gi:1869800.
FEATURES
  source      Location/Qualifiers
             1. .14041
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="10"
             /map="q25"
             /clone="EMBL3"
             /cell_type="HeLa S3"
             /clone_lib="HeLa S3"
             join(11797..11903,12245..>12430)
             /gene="Ki-67"
             11797..11903
             /gene="Ki-67"
             /number=1
             11797..12430
             /gene="Ki-67"
             11904..12244
             /gene="Ki-67"
             /number=1
             12245..>12430
             /gene="Ki-67"
             /number=2
             12335..12430
             /gene="Ki-67"
             /codon_start=1
             /product="monoclonal antibody Ki-67"
             /protein_id="CAA64388.1"
             /db_xref="GI:1869801"
             /db_xref="SWISS-PROT:P46013"
```

```
BASE COUNT   3477 a 3189 c 3327 g 4048 t
ORIGIN

/translation="MMPTRLRLVTIKRSGVDGPHFPLSLSTCLFGR"

Query Match      1.0%; Score 93; DB 9; Length 14041;
Best Local Similarity 100.0%; Pred. No. 8.6e-40;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtggccacagacgagcctgttactacaaagagcggggtcgagcggtccccacctt 60
|||||
Db 12335 ATGTGGCCACAGACGCCCTGGTTACTATCAAAAGGAGCGGGTGCAGGTCGCCCTTT 12394
|||||

QY 61 ccctgagcctcagcacctgctgtttgttggagg 93
|||||
Db 12395 CCCTGAGCCTCAGCACCTGCTGTGTTTGGGAAG 12427
|||||

RESULT 15
LOCUS      HS71L16
DEFINITION Human DNA sequence from clone 71L16 on chromosome Xp11. Contains a
          probable Zinc Finger protein (pseudo)gene, an unknown putative
          gene, a pseudogene with high similarity to part of antigen KI-67, a
          putative Chondroitin 6-Sulfotransferase LIKE gene and a KIAA0267
          LIKE putative Na(+)/H(+) exchanger protein gene. Contains a
          predicted CpG island, ESTs, STSs and GSSs and genomic markers
          DXS1003 and DXS1055, complete sequence.
          AL022165
          AL022165.1 GI:3281985
          HTG; antigen KI-67; CpG island; DXS1003; DXS1055; KIAA0267;
          Sodium/Hydrogen exchanger; Zinc Finger.
          human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 176433)
          Direct Submission
          Submitted (07-JUL-1998) E-mail enquires: humquery@sanger.ac.uk
          Clone requests: clonerequest@sanger.ac.uk
          On Jul 1, 1998 this sequence version replaced gi:3250841.
          During sequence assembly data is compared from overlapping clones.
          Where differences are found these are annotated as variations
          together with a note of the overlapping clone name. Note that the
          variation annotation may not be found in the sequence submission
          corresponding to the overlapping clone, as we submit sequences with
          only a small overlap as described above. This sequence is the
          entire insert of clone 71L16. This sequence has been finished
          according to sequence map criteria as follows. An attempt is made
          to resolve all sequencing problems, such as compressions and
          repeats, but not necessarily within known annotated human repeat
          sequence elements (e.g. Alu). Where the sequence is ambiguous,
          there is an annotation using the 'unsure' feature key.
          This sequence was generated from part of bacterial clone contigs of
          human chromosome X, constructed by the Sanger Centre Chromosome X
          Mapping Group. Further information can be found at
          http://www.sanger.ac.uk/HGP/ChrX
          71L16 is from the library RPC11 constructed at the Roswell Park
          Cancer Institute by the group of Pieter de Jong. For further
          details see http://bacpac.med.buffalo.edu/ VECTOR: pCYPAC2.
FEATURES
  source      Location/Qualifiers
             1. .176433
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="X"
             /map="p11"
             /clone="RP1-71L16"
             /clone_lib="RPC1-1"
             1. .259
             /note="Alu repeat: matches 59. .301 of consensus"
             repeat_region
             271..558
             /note="AluX repeat: matches 1. .289 of consensus"
```

```
repeat_region 675..974
/note="AluX repeat: matches 1..300 of consensus"
repeat_region 1077..1465
/note="L1MB2 repeat: matches 375..772 of consensus"
repeat_region 1469..1767
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 1785..1981
/note="L1MB1 repeat: matches 772..963 of consensus"
repeat_region 2254..2572
/note="AluX repeat: matches 1..301 of consensus"
repeat_region 2649..2931
/note="AluJb repeat: matches 6..299 of consensus"
repeat_region 2942..3254
/note="AluX repeat: matches 1..299 of consensus"
repeat_region 3299..3478
/note="L1MA10 repeat: matches 40..214 of consensus"
repeat_region 3392..3490
/note="L1MB1 repeat: matches 128..225 of consensus"
repeat_region 3551..3730
/note="L1MB2 repeat: matches 326..516 of consensus"
repeat_region 3733..4030
/note="AluJb repeat: matches 1..300 of consensus"
repeat_region complement(4678..4977)
/note="AluX repeat: matches 300..1 of consensus"
repeat_region 5111..5409
/note="AluX repeat: matches 2..301 of consensus"
repeat_region 5414..5745
/note="L1 repeat: matches 3481..3826 of consensus"
repeat_region 5755..6053
/note="AluX repeat: matches 1..303 of consensus"
repeat_region 6083..6373
/note="AluJb repeat: matches 6..297 of consensus"
repeat_region 6404..6651
/note="L1 repeat: matches 3880..4128 of consensus"
repeat_region 6670..6988
/note="AluX repeat: matches 1..303 of consensus"
repeat_region 6990..7615
/note="L1 repeat: matches 4148..4780 of consensus"
repeat_region complement(7634..7926)
/note="AluJb repeat: matches 295..1 of consensus"
repeat_region 8335..8627
/note="AluX repeat: matches 2..299 of consensus"
repeat_region 8631..8923
/note="L1MB1 repeat: matches 41..346 of consensus"
complement(join(<9684..9794,10132..10258,23422..23465,
25850..25960,26720..26768))
/gene="dJ71L16.1"
/note="match: cDNA D31763
match: ESTs AA083688 N41340
match: proteins P51523 Q99676 P21506 Q06730 Q61967 Q62977
Q13401 O14628 O14898 Q03938 Q13359 Q13360 Q15636 Q15928
Q43337 P52738 Q14588 P70591 Q61116 O14913 Q62788 P52737
O15322 Q61491 Q64247 Q62515 P70405
could be a pseudogene"
/product="dJ71L16.1 (probable Zinc Finger protein)"
/evidence="not_experimental
complement(join(9684..9794,10132..10258,23422..23465,
25850..25960,26720..26768))
/gene="dJ71L16.1"
repeat_region complement(10332..10516)
/note="MIR repeat: matches 242..23 of consensus"
10738..11036
/note="AluX repeat: matches 1..299 of consensus"
repeat_region complement(11993..12280)
/note="MER21B repeat: matches 790..499 of consensus"
12284..12453
/note="AluJb repeat: matches 135..302 of consensus"
complement(12471..12954)
/note="MER21B repeat: matches 490..6 of consensus"
13147..13181
/note="L1PA5 repeat: matches 856..890 of consensus"
13234..13536
/note="AluX repeat: matches 1..302 of consensus"

repeat_region 13925..14083
/note="LTR8 repeat: matches 1..146 of consensus"
repeat_region 14101..14326
/note="MER30 repeat: matches 5..229 of consensus"
14632..14815
/note="LTR8 repeat: matches 504..691 of consensus"
14830..15029
/note="AluX repeat: matches 2..201 of consensus"
15180..15219
/note="20 copies 2 mer ta 83% conserved"
complement(15257..15621)
/note="THE1a repeat: matches 364..2 of consensus"
complement(15709..16009)
/note="AluJb repeat: matches 283..1 of consensus"
complement(16088..16389)
/note="AluX repeat: matches 303..1 of consensus"
complement(16494..16794)
/note="AluJb repeat: matches 301..1 of consensus"
17527..17657
/note="AluJb repeat: matches 1..132 of consensus"
17659..17955
/note="AluX repeat: matches 1..299 of consensus"
17957..18125
/note="AluX repeat: matches 134..302 of consensus"
18338..18624
/note="AluX repeat: matches 1..292 of consensus"
19156..19451
/note="AluX repeat: matches 1..296 of consensus"
19959..20250
/note="AluJb repeat: matches 1..292 of consensus"
20253..20553
/note="AluJb repeat: matches 2..302 of consensus"
20832..20929
/note="L1MC3 repeat: matches 458..556 of consensus"
20953..21164
/note="MER20 repeat: matches 1..217 of consensus"
21312..21611
/note="AluY repeat: matches 1..300 of consensus"
22131..22428
/note="AluX repeat: matches 1..298 of consensus"
22429..22730
/note="AluX repeat: matches 1..303 of consensus"
23798..24080
/note="AluJb repeat: matches 21..302 of consensus"
complement(24297..24602)
/note="AluX repeat: matches 303..1 of consensus"
complement(26186..26312)
/note="MIR2 repeat: matches 145..1 of consensus"
26366..26464
/note="MIR repeat: matches 74..177 of consensus"
join(<26877..26976,27815..>28001)
/gene="dJ71L16.2"
/note="match: EST AA322753"
/product="dJ71L16.2 (putative protein)"
/evidence="not_experimental
join(26877..26976,27815..28001)
/gene="dJ71L16.2"
complement(28826..29128)
/note="AluX repeat: matches 297..1 of consensus"
29139..29824
/note="match: ESTs AA487141 AA996220 D20986 AA492440
AA507555 AA486818
match: protein Q14585"
complement(29207..29456)
/note="AluJb repeat: matches 245..1 of consensus"
complement(29916..29985)
/note="L1MA9 repeat: matches 983..910 of consensus"
30031..30328
/note="AluX repeat: matches 1..297 of consensus"
complement(30332..30521)
/note="L1MA8 repeat: matches 777..600 of consensus"
30810..31099
/note="AluY repeat: matches 1..294 of consensus"
```

repeat_region 31758. .32050

Query Match 0.7%; Score 64; DB 9; Length 176433;
Best Local Similarity 100.0%; Pred. No. 2e-23;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8752 ctggccagcttccaagagctctctcaaaacccaggccacactgaggaaactggcaaatcggc 8811
|||||
DB 43604 CTGGCCAGCTTCCAAGAGCTCTCTCAAAACACCCAGGCCACACTGAGGAACCTGGCAAAATGGT 43545
|||||

QY 8812 gctg 8815
|||||

DB 43544 GCTG 43541

Search completed: August 24, 2002, 01:38:22
Job time: 28402 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 1314.37 Seconds
(without alignments)
12756.976 Million cell updates/sec

Title: US-09-700-906A-1_COPY_197_9962
Perfect score: 9766
Sequence: 1 atgtggccacgagacgct.....tcatagggacagtgaagata 9766

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : N_Geneseq_032802.*
- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
 - 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
 - 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
 - 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
 - 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
 - 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
 - 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
 - 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
 - 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
 - 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
 - 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
 - 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
 - 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 - 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 - 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 - 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 - 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 - 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 - 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 - 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 - 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 - 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 - 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
 - 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	9766	100.0	12493	21	AAZ43872	Human cell cycle p
2	9232	94.5	12515	21	AAA64638	cDNA sequence enco
3	8286	84.8	11435	21	AAA64659	cDNA sequence enco
4	629	6.4	733	22	AAH20873	Human Ki-67 protei
5	527	5.4	904	21	AAA64645	Partial sequence M
6	389	4.0	561	21	AAZ80374	Human colon cancer
7	370	3.8	850	21	AAA64643	Partial sequence M
8	344	3.5	861	21	AAA64639	Partial sequence M
9	336	3.4	950	21	AAA64644	Partial sequence M

10	299	3.1	1094	23	AAS69164	DNA encoding novel
c 11	134	1.4	201	19	AAI07117	Human biallelic po
12	110	1.1	158	22	AAF67522	Novel human polynu
13	78	0.8	219	16	AAT20857	Human gene signatu
14	64	0.7	1126	23	AAS69163	DNA encoding novel
15	51	0.5	51	22	AAI27414	Human SNP oligonuc
16	51	0.5	51	22	AAI32696	Human SNP oligonuc
17	51	0.5	51	22	AAI32697	Human SNP oligonuc
18	51	0.5	51	22	AAI32698	Human SNP oligonuc
19	51	0.5	51	22	AAI33289	Human SNP oligonuc
20	51	0.5	51	22	AAI33290	Human SNP oligonuc
21	51	0.5	51	22	AAH79977	Human DNA containi
22	27	0.3	51	22	AAI34120	Human SNP oligonuc
23	26	0.3	53	21	AAA64666	PCR primer used to
24	25	0.3	51	22	AAI27415	Human SNP oligonuc
25	25	0.3	51	22	AAI33291	Human SNP oligonuc
26	24	0.2	24	19	AAI09923	Human biallelic po
c 27	24	0.2	24	19	AAI09349	Human biallelic po
c 28	24	0.2	36	21	AAA64667	PCR primer used to
c 29	23	0.2	23	21	AAZ43869	Human cell cycle p
30	23	0.2	23	21	AAZ43870	Human cell cycle p
31	21	0.2	343	21	AAC29850	Human secreted pro
32	21	0.2	417	22	AAI37373	Human musculoskele
33	21	0.2	417	22	AAI37375	Human musculoskele
34	21	0.2	417	22	AAI37378	Human musculoskele
35	21	0.2	457	22	AAI81035	Human polynucleoti
c 36	21	0.2	489	22	AAI93620	Human polynucleoti
c 37	21	0.2	598	22	AAI60943	Human polynucleoti
c 38	21	0.2	2555	22	AAI59157	Human polynucleoti
c 39	21	0.2	8956	23	ABLI2828	Drosophila melanog
c 40	21	0.2	14388	23	ABLO5670	Drosophila melanog
41	21	0.2	23071	22	AAS26699	Human genomic DNA
c 42	20	0.2	340	17	AAT43951	Sequence flanking
43	20	0.2	422	22	AAF64763	Novel human polynu
44	20	0.2	469	22	ABA43593	Human breast cell
45	20	0.2	469	22	ABA54045	Human foetal liver

ALIGNMENTS

RESULT 1	
AAZ43872	
ID	AAZ43872 standard; DNA; 12493 BP.
XX	
AC	AAZ43872;
DT	10-MAR-2000 (first entry)
XX	
DE	Human cell cycle protein Ki-67 DNA.
XX	
KW	Cell cycle protein; Ki-67; therapy; cell proliferation; allergy;
KW	tumor treatment; autoimmune disease; scar formation; inflammation;
KW	rheumatic disease; transplantation; ds.
XX	
OS	Homo sapiens.
XX	
EH	Key
FT	CDS
FT	Location/Qualifiers
FT	197..9967
FT	/*tag= a
FT	/product= "Ki-67"
XX	
PN	DE19822954-A1.
XX	
PD	25-NOV-1999.
XX	
PF	22-MAY-1998; 98DE-1022954.
XX	
PR	22-MAY-1998; 98DE-1022954.
XX	
FA	(BORS-) FORSCHUNGSZENTRUM BORSTEL ZENT MEDIZIN.
XX	
PI	Flad H, Gerdas J, Boehle A, Deinert I;

XX WPI; 2000-039964/04.
DR P-PSDB; AAY50976.
XX
XX K1-67 gene antisense oligonucleotide -
XX
XX Disclosure; Page 5-19; 36pp; German.
XX
CC This invention describes a novel oligoribo- or oligodeoxyribonucleotide,
CC characterized in that, it hybridizes to mRNA that encodes protein Ki-67
CC at a physiologically acceptable salt concentration. The oligoribo- or
CC oligodeoxyribonucleotide which is complementary to Ki-67, a protein
CC active at all stages of the cell cycle except G₀, is useful for therapy
CC of illnesses with increased cell proliferation and particularly for
CC treatment of tumors, autoimmune diseases, scar formation, inflammation,
CC allergy, rheumatic diseases and defence against transplantation. This
CC sequence encodes the human cell cycle protein Ki-67 which is described in
CC the method of the invention.
XX
SQ Sequence 12493 BP; 4143 A; 3048 C; 2929 G; 2373 T; 0 other;

Query Match 100.0%; Score 9766; DB 21; Length 12493;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9766; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgcccacgacgctgttactatcaaaagagcgggtcgacgtccccacttt 60
Db 197 atgtgcccacgacgctgttactatcaaaagagcgggtcgacgtccccacttt 256
QY 61 ccctcagcctcagcactcgtctgttgggaaggggtattgaaatgacatccgtatccag 120
Db 257 ccctcagcctcagcactcgtctgttgggaaggggtattgaaatgacatccgtatccag 316
QY 121 ctctcgtgtgtccaaacacattgcaagttgaaatccatgacgagggcaatat 180
Db 317 ctctcgtgtgtccaaacacattgcaagttgaaatccatgacgagggcaatat 376
QY 181 cataatttcagttcccaaatcccaacaaagtaattgggtctgtattgatgagcctgta 240
Db 377 cataatttcagttcccaaatcccaacaaagtaattgggtctgtattgatgagcctgta 436
QY 241 cggctaaaaacatggagatgtaataactattattgatcgttccttcagggtatgaaatgaa 300
Db 437 cggctaaaaacatggagatgtaataactattattgatcgttccttcagggtatgaaatgaa 496
QY 301 agtcttcagaatggaaggaaactcaactgaattcccaagaaaaatacgtgaacaggagcca 360
Db 497 agtcttcagaatggaaggaaactcaactgaattcccaagaaaaatacgtgaacaggagcca 556
QY 361 gcacgtcgtgtccaagatctagcttctcttcgacccctgatgagaagctcaagattcc 420
Db 557 gcacgtcgtgtccaagatctagcttctcttcgacccctgatgagaagctcaagattcc 616
QY 421 aaggcctattcaaaaactcactgaaggaagaatttcaggaaaaactcctcaggtacatatcaag 480
Db 617 aaggcctattcaaaaactcactgaaggaagaatttcaggaaaaactcctcaggtacatatcaag 676
QY 481 aatgtcaagaagacagatgacccagatgactcaaaagacagtggtgtcagggaacaact 540
Db 677 aatgtcaagaagacagatgacccagatgactcaaaagacagtggtgtcagggaacaact 736
QY 541 aatgttcattctcagaacatgctggacgttaattggcagaaatgcagctgacccattct 600
Db 737 aatgttcattctcagaacatgctggacgttaattggcagaaatgcagctgacccattct 796
QY 601 ggggattttaaagaatttccacgcttaaatagttagcgcttatggagaattgaaagct 660
Db 797 ggggattttaaagaatttccacgcttaaatagttagcgcttatggagaattgaaagct 856
QY 661 gttcccaactacaaatgtcttggacaatagcaaaaaaatgaaatccctcttgggaagctt 720
Db 857 gttcccaactacaaatgtcttggacaatagcaaaaaaatgaaatccctcttgggaagctt 916

QY 721 tatgagtcagtgaagaaaagagttggatgtaaaatcacaaaaagaaaaatgctctacagtat 780
Db 917 tatgagtcagtgaagaaaagagttggatgtaaaatcacaaaaagaaaaatgctctacagtat 976
QY 781 tgtagaaaaatctggtattacaaaactgattacgcgaacagagaaaagagtgctgtagtcta 840
Db 977 tgtagaaaaatctggtattacaaaactgattacgcgaacagagaaaagagtgctgtagtcta 1036
QY 841 caggaggaaacccaactgttgcctcgcgttaagtcgaagacccaatactggtgagagcgc 900
Db 1037 cagggggagaccacaactgttgcctcgcgttaagtcgaagacccaatactggtgagagcgc 1096
QY 901 cagcgtgtgacagcctgcttccactgaaacaagagcttgaccagacaaggggaagggga 960
Db 1097 cagcgtgtgacagcctgcttccactgaaacaagagcttgaccagacaaggggaagggga 1156
QY 961 agagacgtggagctctgttcaagactccacagcaagcgtgtggtggccagcttctctctat 1020
Db 1157 agagacgtggagctctgttcaagactccacagcaagcgtgtggtggccagcttctctctat 1216
QY 1021 gagccgctaaaaatgaagaccctgtacaatatccacagcaacaaaattctccacaaaaa 1080
Db 1217 gagccgctaaaaatgaagaccctgtacaatatccacagcaacaaaattctccacaaaaa 1276
QY 1081 cataagaacaaaagacctgtatactactggttagaagagaaactgtgaatctgggtataaagt 1140
Db 1277 cataagaacaaaagacctgtatactactggttagaagagaaactgtgaatctgggtataaagt 1336
QY 1141 gaaggctcaagcctggtgataaaaactcttactccacaggaagctttcaactagaaatcga 1200
Db 1337 gaaggctcaagcctggtgataaaaactcttactccacaggaagctttcaactagaaatcga 1396
QY 1201 acaccagctaaaagtgaagatgcagctgcactgcactaaagccagaaaaactctctcc 1360
Db 1397 acaccagctaaaagtgaagatgcagctgcactgcactaaagccagaaaaactctctcc 1456
QY 1261 aaaaaccagaggaagtattcctacagatgtggaagtcttgcctacggaactgaaattcac 1320
Db 1457 aaaaaccagaggaagtattcctacagatgtggaagtcttgcctacggaactgaaattcac 1516
QY 1321 aatgagccatttttaactctgtggtcactcaagtgtgagaggaagatccaaaaaggattcc 1380
Db 1517 aatgagccatttttaactctgtggtcactcaagtgtgagaggaagatccaaaaaggattcc 1576
QY 1381 ctacgaacgctgaagaaattgggcactacagctgcagagatgtgctctgggttaccctggt 1440
Db 1577 ctacgaacgctgaagaaattgggcactacagctgcagagatgtgctctgggttaccctggt 1636
QY 1441 cttagttcagttgatatcaacaaactttggtgattccatttaagttagagtgagggaatacct 1500
Db 1637 cttagttcagttgatatcaacaaactttggtgattccatttaagttagagtgagggaatacct 1696
QY 1501 ttgaaaaagaggcgtgtgctcttgggtggccactaagacctgaactatttggatgaaaac 1560
Db 1697 ttgaaaaagaggcgtgtgctcttgggtggccactaagacctgaactatttggatgaaaac 1756
QY 1561 ttgctctcctaaatcacgctctcaaaaggggagagcccccacccaagaaagagtgctctggtta 1620
Db 1757 ttgctctcctaaatcacgctctcaaaaggggagagcccccacccaagaaagagtgctctggtta 1816
QY 1621 atgcacactccacctgctcctgaagaaaaatcaatcaagggaacagcctcaacccatcaggaaaa 1680
Db 1817 atgcacactccacctgctcctgaagaaaaatcaatcaagggaacagcctcaacccatcaggaaaa 1876
QY 1681 caagagtcaggttccagaataatccatgtggaagtgaaagggcaaaaagcttggtttaaaagcct 1740
Db 1877 caagagtcaggttccagaataatccatgtggaagtgaaagggcaaaaagcttggtttaaaagcct 1936
QY 1741 ccagctccctagtccttaggaaaaactccagttgcccagtgatcaacgcgtaggctctcctcaaa 1800
Db 1937 ccagctccctagtccttaggaaaaactccagttgcccagtgatcaacgcgtaggctctcctcaaa 1996

QY 1801 acagccctgtctccagcagcaaatctcagacagaggttccctaaagagaggagaaaga 1860
Db 1997 acagccctgtctccagcagcaaatctcagacagaggttccctaaagagaggagaaaga 2056
QY 1861 gtggcaactgcttcaaaagagtgctatcagccgaagtcaacatgatattitacag 1920
Db 2057 gtggcaactgcttcaaaagagtgctatcagccgaagtcaacatgatattitacag 2116
QY 1921 atgatatgttccaaaagaagaagtgtcttcggaagcaaatctgatgttgcaaaatca 1980
Db 2117 atgatatgttccaaaagaagaagtgtcttcggaagcaaatctgatgttgcaaaatca 2176
QY 1981 tgggcagatgtagtaaaacttgttgcaaaacaaacacaaaactaaagtataaaatggt 2040
Db 2177 tgggcagatgtagtaaaacttgttgcaaaacaaacacaaaactaaagtataaaatggt 2236
QY 2041 cctcaaaagtcaatgaacaaaaggcaagaagacctgtactccaagaagcctgtgggc 2100
Db 2237 cctcaaaagtcaatgaacaaaaggcaagaagacctgtactccaagaagcctgtgggc 2296
QY 2101 gaagttcacagtccaatttagtacagccacgcaaaactctcctgtaccataataatagg 2160
Db 2297 gaagttcacagtccaatttagtacagccacgcaaaactctcctgtaccataataatagg 2356
QY 2161 aaagctcatactgaaaaagtacatgtcgtctgcacccctacagagtgctcaacaacttc 2220
Db 2357 aaagctcatactgaaaaagtacatgtcgtctgcacccctacagagtgctcaacaacttc 2416
QY 2221 atttccaaacaaaatggacctttaaggaagatctttcaggaaatagctgaaatgtccaag 2280
Db 2417 atttccaaacaaaatggacctttaaggaagatctttcaggaaatagctgaaatgtccaag 2476
QY 2281 acccagtgaaaggagcaaccgagttgacaagcacatgtcacatcgctattttcaaatcca 2340
Db 2477 acccagtgaaaggagcaaccgagttgacaagcacatgtcacatcgctattttcaaatcca 2536
QY 2341 gagaatttgcttggaaaacagtttcaaggaaactgattcaggagaagaaacttgcctccc 2400
Db 2537 gagaatttgcttggaaaacagtttcaaggaaactgattcaggagaagaaacttgcctccc 2596
QY 2401 acctcagagagtttggaggaaatgttcttcagtgacagaatgcagcaaaacagcca 2460
Db 2597 acctcagagagtttggaggaaatgttcttcagtgacagaaatgcagcaaaacagcca 2656
QY 2461 tctgataaatgctctgcaagccctcccttaagcgcgagtgattagagaaaatggaaac 2520
Db 2657 tctgataaatgctctgcaagccctcccttaagcgcgagtgattagagaaaatggaaac 2716
QY 2521 gtacaaaaaccccaggaacacctacaaaaatgacttctctggagacaaaaacttcagat 2580
Db 2717 gtacaaaaaccccaggaacacctacaaaaatgacttctctggagacaaaaacttcagat 2776
QY 2581 actgagacagagccttcaaaaacagtatccactgttaaacaggtcagggaaggtctacagag 2640
Db 2777 actgagacagagccttcaaaaacagtatccactgttaaacaggtcagggaaggtctacagag 2836
QY 2641 ttcaggaaatatacagaagctactctgtgaaaagttaagagtgaagaaacaaatcacgaaatt 2700
Db 2837 ttcaggaaatatacagaagctactctgtgaaaagttaagagtgaagaaacaaatcacgaaatt 2896
QY 2701 gttgagtgatccttaaaaagaggttcagaaggcaacactactacaacaaaggagagaagga 2760
Db 2897 gttgagtgatccttaaaaagaggttcagaaggcaacactactacaacaaaggagagaagga 2956
QY 2761 gagatgaagaaatagaaagaccttttgagacatatagaagaaatattgaattaaaagaa 2820
Db 2957 gagatgaagaaatagaaagaccttttgagacatatagaagaaatattgaattaaaagaa 3016
QY 2821 aacgatgaaaaagatgaaagcaatgaagagatcaagaacttggggcagaaatgtgcacca 2880
Db 3017 aacgatgaaaaagatgaaagcaatgaagagatcaagaacttggggcagaaatgtgcacca 3076
QY 2881 atgtctgacctgacagacctcaagagcttgcttgatcacagaactctcatgaaagacacggca 2940

Db 3077 atgtctgacctgacagacctcaagagcttgcttga tacagaactcatgaaagacacggca 3136
QY 2941 cgtgcccagaaatctctccaaaacccaagatcatgccaagggcaccaaagagtgcagaaagc 3000
Db 3137 cgtgcccagaaatctctccaaaacccaagatcatgccaagggcaccaaagagtgcagaaagc 3196
QY 3001 aaaaactactaaaaatgccttgcctgcccagtcattacaacccagaaaccaaataaacccccacacac 3060
Db 3197 aaaaactactaaaaatgccttgcctgcccagtcattacaacccagaaaccaaataaacccccacacac 3256
QY 3061 acaaaaacacagttgaaggcatcccttggggaaaagttaggtgtgaaaaagaagctccttagca 3120
Db 3257 acaaaaacacagttgaaggcatcccttggggaaaagttaggtgtgaaaaagaagctccttagca 3316
QY 3121 gtccgcaagtccacacggagctcaggggagaccacgacacacacacagagagccagcagga 3180
Db 3317 gtccgcaagtccacacggagctcaggggagaccacgacacacacagagagccagcagga 3376
QY 3181 gatgcaagagcatcagaacgctttaaggagcttccaaaagcagatccttggaccacagcgc 3240
Db 3377 gatgcaagagcatcagaacgctttaaggagcttccaaaagcagatccttggaccacagcgc 3436
QY 3241 cgtgttaacttggaaatgaagaagtggccaaagacgccttaaggaaggccacagtcactagaa 3300
Db 3437 cgtgttaacttggaaatgaagaagtggccaaagacgccttaaggaaggccacagtcactagaa 3496
QY 3301 gacctggctggccttcaaaagagcttctccagacaccaggtccctctgagaaatcaatgact 3360
Db 3497 gacctggctggccttcaaaagagcttctccagacaccaggtccctctgagaaatcaatgact 3556
QY 3361 gatgaaaaacttaccaaaatagcctgcgaaatctccaccaccagaatcagtggaactcca 3420
Db 3557 gatgaaaaacttaccaaaatagcctgcgaaatctccaccaccagaatcagtggaactcca 3616
QY 3421 acaagcacaacaaatggccttaagagaagtctcaggaagagcagatgtaggaagaattc 3480
Db 3617 acaagcacaacaaatggccttaagagaagtctcaggaagagcagatgtaggaagaattc 3676
QY 3481 tttagcactcaggaactcaacacatcagcaggaagagccatctttagcggcgaacccagca 3540
Db 3677 tttagcactcaggaactcaacacatcagcaggaagagccatctttagcggcgaacccagca 3736
QY 3541 ggaggtgtagaagaagacattaaagcatttatgggaactccagtgagaaaacttggacctg 3600
Db 3737 ggaggtgtagaagaagacattaaagcatttatgggaactccagtgagaaaacttggacctg 3796
QY 3601 gcaggaaactttaccttggcggcaaaagacagctcacagactccttaaggaagggcccaggct 3660
Db 3797 gcaggaaactttaccttggcggcaaaagacagctcacagactccttaaggaagggcccaggct 3856
QY 3661 ctagaagacctggctggcctttaaagagcttctccagactccttggcggcgaacccaggaattc 3720
Db 3857 ctagaagacctggctggcctttaaagagcttctccagactccttggcggcgaacccaggaattc 3916
QY 3721 gtggctgctgtgttaaaacccactaaaaatccctgcgactctccacagtcagaccagtgagac 3780
Db 3917 gtggctgctgtgttaaaacccactaaaaatccctgcgactctccacagtcagaccagtgagac 3976
QY 3781 accccaacaaagcacaaaagcaacgcccagagaagtatcaggaaaacagatgtagagggga 3840
Db 3977 accccaacaaagcacaaaagcaacgcccagagaagtatcaggaaaacagatgtagagggga 4036
QY 3841 gaactcttagcgtgcaggaaatctaattgcccattcagcagggcaaaagcccatgcacgcgcctaaa 3900
Db 4037 gaactcttagcgtgcaggaaatctaattgcccattcagcagggcaaaagcccatgcacgcgcctaaa 4096
QY 3901 ccatcagtaggtgaaagagaagacatcataatttgggggaactccagtcgcaagaactg 3960
Db 4097 ccatcagtaggtgaaagagaagacatcataatttgggggaactccagtcgcaagaactg 4156
QY 3961 gaactgacagagaacttaaccggcagcaagagacggcccaaaactccttaagggaagggcc 4020

Db 4157 gacctgacagaaacttaacggcagcaagacagcggccacaaaactcctaaggaaagccc 4216
QY 4021 caggctctgaaagacctgactggctttaagagagctcttccagaccctggctcatactgaa 4080
Db 4217 caggctctgaaagacctgactggctttaagagagctcttccagaccctggctcatactgaa 4276
QY 4081 gaagcagtgctctgctgcaaaactactataaatagccttcgcgaattcttccaccagaaatca 4140
Db 4277 gaagcagtgctctgctgcaaaactactataaatagccttcgcgaattcttccaccagaaatca 4336
QY 4141 gcagacaccccaacaagacacaagaagcagcccaagacacacttttgagaagaaggacgta 4200
Db 4337 gcgaqcccccacaagacacaagaagcagcccaagacacttttgagaagaaggacgta 4396
QY 4201 cagaagagctctcagccctgaagaagctcacacagacatcaggggaaaccacacacaca 4260
Db 4397 cagaagagctctcagccctgaagaagctcacacagacatcaggggaaaccacacacaca 4456
QY 4261 gataaagtcaccagaggtgaggataaagcatcaacgcgttttagggaaaactgcaaaaacag 4320
Db 4457 gataaagtcaccagaggtgaggataaagcatcaacgcgttttagggaaaactgcaaaaacag 4516
QY 4321 aaactggaccagcagcaagtgtaactggtagcaagagcagcccaaaaactaaggaaaag 4380
Db 4517 aaactggaccagcagcaagtgtaactggttagcaagagcagcccaaaaactaaggaaaag 4576
QY 4381 gccaaaccttagaagacctggctggctggaagagctcttccagacacacagtatgcact 4440
Db 4577 gccaaaccttagaagacctggctggctggaagagctcttccagacacacagtatgcact 4636
QY 4441 gacaagcccacgactcaggaataactaccaaataagcctgcagatcacacaccagaccaca 4500
Db 4637 gacaagcccacgactcaggaataactaccaaataagcctgcagatcacacaccagaccaca 4696
QY 4501 gtgagacaccaaagctccaagccacagctccaagagaagctcaggaaagtggacgta 4560
Db 4697 gtgagacaccaaagctccaagccacagctccaagagaagctcaggaaagtggacgta 4756
QY 4561 gaagaagaattcttcgactcaggaagacgaacaccatcagcaggcaagccatgcacaca 4620
Db 4757 gaagaagaattcttcgactcaggaagacgaacaccatcagcaggcaagccatgcacaca 4816
QY 4621 ccacaaaccagcagtaagtggtggaataaacatctacgcatttatgggaactccagtcgag 4680
Db 4817 ccacaaaccagcagtaagtggtggaataaacatctacgcatttatgggaactccagtcgag 4876
QY 4681 aaactggacctgcagagaacttaactggcagcaagacgcgtacaaaactcctaaggaa 4740
Db 4877 aaactggacctgcagagaacttaactggcagcaagacgcgtacaaaactcctaaggaa 4936
QY 4741 aagggccaggctctagaagacctggctgtttaaagagctcttccagacacgcaggtcac 4800
Db 4937 aagggccaggctctagaagacctggctgtttaaagagctcttccagacacgcaggtcac 4996
QY 4801 actgaggaataactagactaaactgaaactgccaagtgagcctgcgaataatttcaacaacca 4860
Db 4997 actgaggaataactagactaaactgaaactgccaagtgagcctgcgaataatttcaacaacca 5056
QY 4861 gacctagacaaaacccagcaagctccaagcagcgtccaagacatccctgggaaagtg 4920
Db 5057 gacctagacaaaacccagcaagctccaagcagcgtccaagacatccctgggaaagtg 5116
QY 4921 ggcgtgaaagaagagctccttagcagttggcaagctcacacagacatcaggagagactaca 4980
Db 5117 ggcgtgaaagaagagctccttagcagttggcaagctcacacagacatcaggagagactaca 5176
QY 4981 cacacacacagagccaaacagagatggttaagagcatgaaagcatttatggagttccca 5040
Db 5177 cacacacacagagccaaacagagatggttaagagcatgaaagcatttatggagttccca 5236
QY 5041 aagcagatcttagcagcagcaagcttaactggcagcaagagcagctgagaactcct 5100
Db 5237 aagcagatcttagcagcagcaagcttaactggcagcaagagcagctgagaactcct 5296

QY 5101 aagggaaagtctgaagtccctgaagacctggccgggttctcatcgagctcttccagacacca 5160
Db 5297 aagggaaagtctgaagtccctgaagacctggccgggttctcatcgagctcttccagacacca 5356
QY 5161 agtcacactaagaagaatcaatgactaaatgaaaaactaccaaagtatctctacagagcttca 5220
Db 5357 agtcacactaagaagaatcaatgactaaatgaaaaactaccaaagtatctctacagagcttca 5416
QY 5221 cagccagacctagtgagacaccccaacagctccaagccacagcccaagaagaagtctcagg 5280
Db 5417 cagccagacctagtgagacaccccaacagctccaagccacagcccaagaagaagtctcagg 5476
QY 5281 aagcgagacactgaagaagaatttttagcattttaggaacaaacgcgcctcagcaggcaaa 5340
Db 5477 aagcgagacactgaagaagaatttttagcattttaggaacaaacgcgcctcagcaggcaaa 5536
QY 5341 gccatgcacacaccccaacccagcagtaggtggaagagaagaacatcaacgctttttggga 5400
Db 5537 gccatgcacacaccccaacccagcagtaggtggaagagaagaacatcaacgctttttggga 5596
QY 5401 actccagtgcaagaactggacagccagcaggaattttacctggcagcaatagacggctacaa 5460
Db 5597 actccagtgcaagaactggacagccagcaggaattttacctggcagcaatagacggctacaa 5556
QY 5461 actcgtlaaggaaagcccgagctcttagaagaactgactggcttcagagagctttttccag 5520
Db 5657 actcgtlaaggaaagcccgagctcttagaagaactgactggcttcagagagctttttccag 5716
QY 5521 acacctgcaactgataccccacagctgataagaactaccacaaaaaactctctgcacaa 5580
Db 5717 acacctgcaactgataccccacagctgataagaactaccacaaaaaactctctgcacaa 5776
QY 5581 tctcccaatacgaagccagcggagaccccaacacacacaaagcaacggcccaagagaagc 5640
Db 5777 tctcccaatacgaagccagcggagaccccaacacacacaaagcaacggcccaagagaagc 5836
QY 5641 ctcaagaaagcagactagaggagaatttttagcatctcaggaaacttaacacatcagca 5700
Db 5837 ctcaagaaagcagactagaggagaatttttagcatctcaggaaacttaacacatcagca 5896
QY 5701 ggcaagccatgcacacgcctaaagcagcagtaggtggaagagaagaacatcaacacattt 5760
Db 5897 ggcaagccatgcacacgcctaaagcagcagtaggtggaagagaagaacatcaacacattt 5956
QY 5761 gtgggactccagtggagaaactggcctgcctaggaatttacctggcagcagagaagcgg 5820
Db 5957 gtgggactccagtggagaaactggcctgcctaggaatttacctggcagcagagaagcgg 6016
QY 5821 ccacaaactcctaaagaaagggccaaaggctctagaagatctggctggcttcaaaagagctc 5880
Db 6017 ccacaaactcctaaagaaagggccaaaggctctagaagatctggctggcttcaaaagagctc 6076
QY 5881 ttccagacacccaggtcacactgagggaatacaatgcacgatagaacaaatcacaagaatctcc 5940
Db 6077 ttccagacacccaggtcacactgagggaatacaatgcacgatagaacaaatcacaagaatctcc 6136
QY 5941 tgcaaatctccacacacagacccagtcacaaaccccaacagctcccaagcaacgactcaag 6000
Db 6137 tgcaaatctccacacacagacccagtcacaaaccccaacagctcccaagcaacgactcaag 6196
QY 6001 atatccttgggaaagttaggtgtgaagaagaggttctccagctcggaagctcacacag 6060
Db 6197 atatccttgggaaagttaggtgtgaagaagaggttctccagctcggaagctcacacag 6256
QY 6061 acgtcaggaagaaccacacagacacacagagagacagcgggagatggaaagagcatcaaaa 6120
Db 6257 acgtcaggaagaaccacacagacacacagagagacagcgggagatggaaagagcatcaaaa 6316
QY 6121 gcggttaaggaattctcaaaagcagatgcctggaccacagcaaaactatggaactggagtgag 6180
Db 6317 gcggttaaggaattctcaaaagcagatgcctggaccacagcaaaactatggaactggagtgag 6376

Db 8537 actggcagcaggagagcgccaaagcaccagggaaagtgtcccaagccatagaaaccta 8596
QY 8401 gctggcttcaaaagacccagcagcaggttcacactgaagaatacaatgactgatgacaaaacc 8460
Db 8597 gctggcttcaaaagacccagcagcaggttcacactgaagaatacaatgactgatgacaaaacc 8556
QY 8461 actaaataacctgcaaatcatcaccagnactagaagacacccgcaacaagctcaaaagaga 8520
Db 8657 actaaataacctgcaaatcatcaccagaactagaagacacccgcaacaagctcaaaagaga 8716
QY 8521 cggccagagacagtcgccagaaagttagaagtgaaggaggagctgttagcagttggcaag 8580
Db 8717 cggccagagacagtcgccagaaagttagaagtgaaggaggagctgttagcagttggcaag 8776
QY 8581 ctacacaaacctcaggggagacacgcacaccgcgacaagagcggtagtgagggcaaa 8640
Db 8777 ctacacaaacctcaggggagacacgcacaccgcgacaagagcggtagtgagggcaaa 8836
QY 8641 ggcacgaagcatttaagcaacctgcaaaagcgaaacgtggagcagaagaatgtaattggc 8700
Db 8837 ggcacgaagcatttaagcaacctgcaaaagcgaaacgtggagcagaagaatgtaattggc 8896
QY 8701 agcagagacagccaagcaccacctaaggaagggcccaacccctggaaagacctggccagc 8760
Db 8897 agcagagacagccaagcaccacctaaggaagggcccaacccctggaaagacctggccagc 8956
QY 8761 ttccaagagctctcaaacaccagccacactgaggaactggcgaatgtgtgctgctgat 8820
Db 8957 ttccaagagctctcaaacaccagccacactgaggaactggcgaatgtgtgctgctgat 9016
QY 8821 agctttacaagcgtccaaagcgaacacactgacagtggaaacacctctaaaaatatccaga 8880
Db 9017 agctttacaagcgtccaaagcgaacacactgacagtggaaacacctctaaaaatatccaga 9076
QY 8881 agagttcttcggccctaaagttagaacccgtggagacgtggttaagcaccagagacct 8940
Db 9077 agagttcttcggccctaaagttagaacccgtggagacgtggttaagcaccagagacct 9136
QY 8941 gtaaatcacaagcaaaagcaaacacttcctgtcccccactgccttcagaggggaggt 9000
Db 9137 gtaaatcacaagcaaaagcaaacacttcctgtcccccactgccttcagaggggaggt 9196
QY 9001 ggcacaagatggaagcgtcacgggaaccaaagaggtcgctgcatgcagcacccagaggaa 9060
Db 9197 ggcacaagatggaagcgtcacgggaaccaaagaggtcgctgcatgcagcacccagaggaa 9256
QY 9061 attgtgagagagctgccagccagcaagaagcagaggttgctccccaaggcaagagcaaa 9120
Db 9257 attgtgagagagctgccagccagcaagaagcagaggttgctccccaaggcaagagcaaa 9316
QY 9121 tcattccgaacccgtggtcatcatgaagagaagtttgagagactctgcgaagaagaattgaa 9180
Db 9317 tcattccgaacccgtggtcatcatgaagagaagtttgagagactctgcgaagaagaattgaa 9376
QY 9181 cctgcggaagagctgaacgcgaacgcacatgaaacccaacaaagaggaacacaaattacaa 9240
Db 9377 cctgcggaagagctgaacgcgaacgcacatgaaacccaacaaagaggaacacaaattacaa 9436
QY 9241 gactcgtccctgaaataaagggaatatccctgcgtccagcagccaaagataagactgag 9300
Db 9437 gactcgtccctgaaataaagggaatatccctgcgtccagcagccaaagataagactgag 9496
QY 9301 gcagaacagcaataaactgaggtcttctgtattagcagaagaatagaaataaacagaaat 9360
Db 9497 gcagaacagcaataaactgaggtcttctgtattagcagaagaatagaaataaacagaaat 9556
QY 9361 gaaagaagccatgaagaccctcccagagatggacattcagaatccagatgatggagcc 9420
Db 9557 gaaagaagccatgaagaccctcccagagatggacattcagaatccagatgatggagcc 9616
QY 9421 cggaaacccatccttagagacaaaagtccactgagacaaaaggtgcttaggtctgctaga 9480
Db 9617 cggaaacccatccttagagacaaaagtcactgagacaaaaggtgcttaggtctgctaga 9676

QY 9481 cagaatgagagctcccagcctaaagtgtgagagagagcgagggcagaagaagtgcgaag 9540
Db 9677 cagaatgagagctcccagcctaaagtgtgagagagagcgagggcagaagaagtgcgaag 9736
QY 9541 gttctcatgcagaatcagaagaagggaagcagcagggaattcagactccatgtgcctg 9600
Db 9737 gttctcatgcagaatcagaagaagggaagcagcagggaattcagactccatgtgcctg 9796
QY 9601 agatcaagaagaacaaaaagccagcctgcagcaagcacitttgagagcaaatctgtgcag 9660
Db 9797 adatcaagaagaacaaaaagccagcctgcagcaagcacitttgagagcaaatctgtgcag 9856
QY 9661 agagtcaacgcgagtgctcaagaggtgtgcagaaaaatccaaaagagctgaggacaattgtg 9720
Db 9857 agagtcaacgcgagtgctcaagaggtgtgcagaaaaatccaaaagagctgaggacaattgtg 9916
QY 9721 tgtgtcaagaataatacaaccagaaagtcataagggacagtgaagata 9766
Db 9917 tgtgtcaagaataatacaaccagaaagtcataagggacagtgaagata 9962

RESULT 2
AAA64658
ID AAA64658 standard; cDNA; 12515 BP.
XX
AC AAA64658;
XX
DT 02-JAN-2001 (first entry)
XX
DE cDNA sequence encoding the antigen of monoclonal antibody Ki-67.
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW antibody Ki-67; Ss.
OS Homo sapiens.
XX
XX WO2000050595-A2.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-US04929.
XX
XX 25-FEB-1999; 99US-0257417.
XX
XX (GOUT/) GOUT I.
XX (RODN/) RODNIN N.
XX (FILO/) FILONENKO V.
XX (MATS/) MATSUKA G.
XX (SCAN/) SCANLAN M.
XX (OLDL/) OLD L.
XX (BILY/) BILYNSKY B.
XX
XX Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
PI
XX WPI; 2000-572092/53.
XX
XX Novel isolated nucleic acid molecules for diagnosing and treating -
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
XX
XX Claim 17; Page 77-81; 94pp; English.
XX
XX The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a polynucleotide of the invention.
XX
XX Sequence 12515 BP; 4164 A; 3050 C; 2928 G; 2373 T; 0 other;

Query Match 94.5%; Score 9232; DB 21; Length 12515;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 9762; Conservative 0; Mismatches 1; Indels 6; Gaps 2;

Qy	1	atgtggccacgagcgcctggttactatcaaaaggagcgggtgcagcgtcccccattt	60
Db	197	atgtggccacgagcgcctggttactatcaaaaggagcgggtgcagcgtcccccattt	256
Qy	61	ccccgagccctcagcacctgcttggtaagggtattgaattgacatccgtatccag	120
Db	257	ccccgagccctcagcacctgcttggtaagggtattgaattgacatccgtatccag	316
Qy	121	cttcctgtgtgtcaaaacacattgcaaaagtgaattcccatgagcagaggaatatta	180
Db	317	cttcctgtgtgtcaaaacacattgcaaaagtgaattcccatgagcagaggaatatta	376
Qy	181	cataatttcagttccacaaatccacacaaagtaaatgggtgttattgatgagcctgta	240
Db	377	cataatttcagttccacaaatccacacaaagtaaatgggtgttattgatgagcctgta	436
Qy	241	cggctaaaacatggagatgtaataactattatgtatccttcagatgataaataaa	300
Db	437	cggctaaaacatggagatgtaataactattatgtatccttcagatgataaataaa	496
Qy	301	agttctcagaatggaggaagtcaactgaatttccaaagaaaaatacgtgaacagagcca	360
Db	497	agttctcagaatggaggaagtcaactgaatttccaaagaaaaatacgtgaacagagcca	556
Qy	361	gcacgtcgtgtctcaagatctagcttctctctgtgacctgatgagaagctcaagattcc	420
Db	557	gcacgtcgtgtctcaagatctagcttctctctgtgacctgatgagaagctcaagattcc	616
Qy	421	aaggcctattcaaaatcacgaagaaaagtttcagaaaatcctcaggtacatatcaag	480
Db	617	aaggcctattcaaaatcacgaagaaaagtttcagaaaatcctcaggtacatatcaag	676
Qy	481	aatgtcaagaagacagtaccgcagatgactcaaaagacagtggtgctcagggaacaact	540
Db	677	aatgtcaagaagacagtaccgcagatgactcaaaagacagtggtgctcagggaacaact	736
Qy	541	aatgttcatctctcagaacatgctggaagtaatgagcagaataatgcagctgatccccattct	600
Db	737	aatgttcatctctcagaacatgctggaagtaatgagcagaataatgcagctgatccccattct	796
Qy	601	ggggattttaagaaatttcagcgttaaatagtgagccgttatggagaattgagct	660
Db	797	ggggattttaagaaatttcagcgttaaatagtgagccgttatggagaattgagct	856
Qy	661	gtccacctacacaatgtcttgacaatagcaaaaaaatgaaatcccttttggagctt	720
Db	857	gtccacctacacaatgtcttgacaatagcaaaaaaatgaaatcccttttggagctt	916
Qy	721	tatgagtcagtgagaagaaggttgatgtataaatcacaaaaagaaaaatgtctcagat	780
Db	917	tatgagtcagtgagaagaaggttgatgtataaatcacaaaaagaaaaatgtctcagat	976
Qy	781	tgtagaaaatctggattacaactgattacgcaacagagagaagaagtgtgatgttta	840
Db	977	tgtagaaaatctggattacaactgattacgcaacagagagaagaagtgtgatgttta	1036
Qy	841	caggggagagacccaactgttgctcgcgtgaagtcaagaccaaatactggggagcgc	900
Db	1037	caggggagagacccaactgttgctcgcgtgaagtcaagaccaaatactggggagcgc	1096
Qy	901	cacgtctggcagagcctgtctcaacctgaacaagagcttgaccagacaaggggagggga	960
Db	1097	cacgtctggcagagcctgtctcaacctgaacaagagcttgaccagacaaggggagggga	1156
Qy	961	agagacgtggagctgtgtcagactcccagcagcgtgtgggcccagctttcctctctat	1020
Db	1157	agagacgtggagctgtgtcagactcccagcagcgtgtgggcccagctttcctctctat	1216
Qy	1021	gagcgggtaaaaatgaagaccctgtacaataattcacagcaacaaaattctccaaaaa	1080
Db	1217	gagcgggtaaaaatgaagaccctgtacaataattcacagcaacaaaattctccaaaaa	1276
Qy	1081	cataagaacaaagacctgtatactactggtagaagaatctgtgaatctgggtaaaaagt	1140
Db	1277	cataagaacaaagacctgtatactactggtagaagaatctgtgaatctgggtaaaaagt	1336
Qy	1141	gaagcttcaagcctgggtgataaaaactctactcccagagaagctttcaactagaaatcga	1200
Db	1337	gaagccttcaagcctgggtgataaaaactctactcccagagaagctttcaactagaaatcga	1396
Qy	1201	acaccagctaaagtgtgaagatgcagctgactgcactaaagccagaaaaatctctctcc	1260
Db	1397	acaccagctaaagtgtgaagatgcagctgactgcactaaagccagaaaaatctctctcc	1456
Qy	1261	aaaaaccagaggaagtattctctacagatgtggaagtctcctacggaaaaactgaaattcac	1320
Db	1457	aaaaaccagaggaagtattctctacagatgtggaagtctcctacggaaaaactgaaattcac	1516
Qy	1321	aatgagccatttttaactctgtggtcactcaagtgtgagaggaagatccaaaagattcc	1380
Db	1517	aatgagccatttttaactctgtggtcactcaagtgtgagaggaagatccaaaagattcc	1576
Qy	1381	ctcagcaagcctggaataattgggcactacagctgacagatgtctctgggttacctggt	1440
Db	1577	ctcagcaagcctggaataattgggcactacagctgacagatgtctctgggttacctggt	1636
Qy	1441	cttagttcagttgatatacaaaactttggtgattccattaatgagagtgaggaatacct	1500
Db	1637	cttagttcagttgatatacaaaactttggtgattccattaatgagagtgaggaatacct	1696
Qy	1501	ttgaaaaagagcgtgtgtcctttggtggccactaaagccctgaactatttcatgaaaaac	1560
Db	1697	ttgaaaaagagcgtgtgtcctttggtggccactaaagccctgaactatttcatgaaaaac	1756
Qy	1561	ttgcctctctaatcacgcctctcaaaaggggagagcccccacaaaagaagctctctggtta	1620
Db	1757	ttgcctctctaatcacgcctctcaaaaggggagagcccccacaaaagaagctctctggtta	1816
Qy	1621	atgcacactccacctgtccttgagaagaaaaatcatcaagggaacagcctcaaacatcaggaaaa	1680
Db	1817	atgcacactccacctgtccttgagaagaaaaatcatcaagggaacagcctcaaacatcaggaaaa	1876
Qy	1681	caagagtcagggttcagaaaaatccatgtggaagtgaaggcacaagcttggttataagccct	1740
Db	1877	caagagtcagggttcagaaaaatccatgtggaagtgaaggcacaagcttggttataagccct	1936
Qy	1741	ccagctcctagctctagaaaaactccagttgccagtgaataacgcctgaggtctctgcaaa	1800
Db	1937	ccagctcctagctctagaaaaactccagttgccagtgaataacgcctgaggtctctgcaaa	1996
Qy	1801	acagccctctgtctcagcagcaaaaatctcagacagaggttccctaaagagagagagaaaaa	1860
Db	1997	acagccctctgtctcagcagcaaaaatctcagacagaggttccctaaagagagagagaaaaa	2056
Qy	1861	gtggcaacctgctctcaaaaagagtgctctatcagccgaagtcaacatgattttcacag	1920
Db	2057	gtggcaacctgctctcaaaaagagtgctctatcagccgaagtcaacatgattttcacag	2116
Qy	1921	atgatattgtccaaaagaagaagtgtgcttcgggaagcaaaatctgattgtgcaaaatca	1980
Db	2117	atgatattgtccaaaagaagaagtgtgcttcgggaagcaaaatctgattgtgcaaaatca	2176
Qy	1981	tgggcagatgtagtaaaaacttggtcaaaaacaaacacaaactaaagtataaaaatggt	2040
Db	2177	tgggcagatgtagtaaaaacttggtcaaaaacaaacacaaactaaagtataaaaatggt	2236
Qy	2041	ctcctaaaggtcaatgaacaaaagcacaagacacctgtactccaaaagaagcctgtgggc	2100
Db	2237	ctcctaaaggtcaatgaacaaaagcacaagacacctgtactccaaaagaagcctgtgggc	2296
Qy	2101	gaagttcacagtcaatttagtacagggccacgcaaacctctcttgtaccataataaggg	2160

Db	2297		gaagttcaacagtcacatttagtacaggccacgcgcaaatctcccttgytaccataataataggg	2356
QY	2161		aaagctcactactgaaaaagtacatgtgcctgtcgacctacagagtgctccaacaacttc	2220
Db	2357		aaagctcactactgaaaaagtacatgtgcctgtcgacctacagagtgctccaacaacttc	2416
QY	2221		attccaacaaaaaatggaactttaaggaaagatctttcagggaatagctgaaatggtccaag	2280
Db	2417		attccaacaaaaaatggaactttaaggaaagatctttcagggaatagctgaaatggtccaag	2476
QY	2281		accagtgaaagagcaaccgcagttgcacaagcacatgctcacatcgctatttccaaattcca	2340
Db	2477		accocagtgaaagagcaaccgcagtgacaagcacatgctcacatcgctatttccaaattcca	2536
QY	2341		gagaatttgccttggaacacagttctcaggaactgattcaggagaagaactctgctcccc	2400
Db	2537		gagaatttgccttggaacacagttctcaggaactgattcaggagaagaactctgctcccc	2596
QY	2401		acctcagagagttttggaggaaatgtgttcttcagtgccacagaatgcagcaaaacagcca	2460
Db	2597		acctcagagagttttggaggaaatgtgttcttcagtgccacagaatgcagcaaaacagcca	2656
QY	2461		tcgtataaatgctctgcaagccctcccttaagacggcagtgatttagagaaaaatgaaac	2520
Db	2657		tcgtataaatgctctgcaagccctcccttaagacggcagtgatttagagaaaaatgaaac	2716
QY	2521		gtagcaaaacgccaggaacactcacaatgactctcttgagacaaacactcttcagat	2580
Db	2717		gtagcaaaacgccaggaacactcacaatgactctcttgagacaaacactcttcagat	2776
QY	2581		actgagacagagccttcaaaaacagtatccactgttaaacaggtcaggaaaggtctacagag	2640
Db	2777		actgagacagagccttcaaaaacagtatccactgttaaacaggtcaggaaaggtctacagag	2836
QY	2641		ttcaggaataacagaaagctactctgtgaaagttaaagtgaaagaaacaaatacagaaatt	2700
Db	2837		ttcaggaataacagaaagctactctgtgaaagttaaagtgaaagaaacaaatacagaaatt	2896
QY	2701		gttgagtgcatcctctaaaaagaggtcgagaaggcaacactactacacaagaaggagagaagga	2760
Db	2897		gttgagtgcatcctctaaaaagaggtcgagaaggcaacactactacacaagaaggagagaagga	2956
QY	2761		gagatgaagaaatagaaagaccttttgagacataaaaggaaaaatttgaattaaagaagaa	2820
Db	2957		gagatgaagaaatagaaagaccttttgagacataaaaggaaaaatttgaattaaagaagaa	3016
QY	2821		aacgatgaaagatgaaagcaatgaaagatcaagaaacttggggcgagaatgtgcacca	2880
Db	3017		aacgatgaaagatgaaagcaatgaaagatcaagaaacttggggcgagaatgtgcacca	3076
QY	2881		atgtctgacctgacagacctcaagagcttgcctgtatcacagaactctatgaaagacacggca	2940
Db	3077		atgtctgacctgacagacctcaagagcttgcctgtatcacagaactctatgaaagacacggca	3136
QY	2941		ctgycgagaatctctccaaaccccaagaatcatgccaaggacccaagagtgagaaagggc	3000
Db	3137		ctgycgagaatctctccaaaccccaagaatcatgccaaggacccaagagtgagaaagggc	3196
QY	3001		aaatcactaaaaatgccctgccagtcattacaacgagaaccaataaaccacacacac	3060
Db	3197		aaatcactaaaaatgccctgccagtcattacaacgagaaccaataaaccacacacacac	3256
QY	3061		acaaaacacagttgaaggcatccctcggggaaagttaggtgtgaaagaaagagctcccttagca	3120
Db	3257		acaaaacacagttgaaggcatccctcggggaaagttaggtgtgaaagaaagagctcccttagca	3316
QY	3121		gtcggcaagtccacacgacgtcaggggagaccacgcacacgcacagagagccagcagga	3180
Db	3317		gtcggcaagtccacacgacgtcaggggagaccacgcacacgcacagagagccagcagga	3376
QY	3181		gatggcaagagcatcagaacgtttaagaggtctccaaagcadatcttgaccacagagcc	3240
Db				

Db	3377		gatggcaagagcatcgaacogttaaaggagttcccaaaagcagatccttgagaccacagagcc	3436
QY	3241		cggtgaactggaatgaagaagtggccaagaacgccttaaggaaagagggcccgagtcactagaa	3300
Db	3437		cggtgaactggaatgaagaagtggccaagaacgccttaaggaaagagggcccgagtcactagaa	3496
QY	3301		gaactggctggcttcaaaagagctctccagacaccagggtccctctgagggaatcaatgact	3360
Db	3497		gaactggctggcttcaaaagagctctccagacaccagggtccctctgagggaatcaatgact	3556
QY	3361		gatgagaaaactaccaaaatagcctgcaaatct--ccaccacagaatacagtgggacact	3417
Db	3557		gatgagaaaactaccaaaatagcctgcaaatctcccccaccacagaatacagtgggacact	3616
QY	3418		ccaacaagcacaaagcaatggcctaagagaagtctcaggaaaacagatgtagaggaagaa	3477
Db	3617		ccaacaagcac--gcaatggcctaagagaagtctcaggaaaacagatgtagaggaagaa	3673
QY	3478		ttcttagcactcaggaaaactaacaccatcagcaggggaaagccatgtcttacgcgccaaacca	3537
Db	3674		ttcttagcactcaggaaaactaacaccatcagcaggggaaagccatgtcttacgcgccaaacca	3733
QY	3538		gcaggaggtgatgagaaagacattaaagcatttatgggaactccagtgcaagaactggac	3597
Db	3734		gcaggaggtgatgagaaagacattaaagcatttatgggaactccagtgcaagaactggac	3793
QY	3598		ctggcaggaaactttacctggcagcaaaagacagctacagactcctaaagaaaagggccacg	3657
Db	3794		ctggcaggaaactttacctggcagcaaaagacagctacagactcctaaagaaaagggccacg	3853
QY	3658		gctctagaagaactggctggctttaagagctcttccagactcctgggtcacaccggaggaa	3717
Db	3854		gctctagaagaactggctggctttaagagctcttccagactcctgggtcacaccggaggaa	3913
QY	3718		ttagtgtctgctgggttaaaaccactaaaataccctcgagctctccacagtcagaccagtg	3777
Db	3914		ttagtgtctgctgggttaaaaccactaaaataccctcgagctctccacagtcagaccagtg	3873
QY	3778		gacaccccaacagcaacaagcaacgacccaagagaglatcaggaagcagatgtagag	3837
Db	3974		gacaccccaacagcaacaagcaacgacccaagagaglatcaggaagcagatgtagag	4033
QY	3838		ggagaactcttagcgtgcagggaatcttaatgcactcagcgggcaaaagccatgcacacgcct	3897
Db	4034		ggagaactcttagcgtgcagggaatcttaatgcactcagcgggcaaaagccatgcacacgcct	4093
QY	3898		aaacctcagtaggtgaagagaaaagcatcatatttctgggaactccagtcgcagaa	3957
Db	4094		aaacctcagtaggtgaagagaaaagcatcatatttctgggaactccagtcgcagaa	4153
QY	3958		ctggacctgacagagaaacttaaccgagcagagagcggccacaacactcctaaagggaag	4017
Db	4154		ctggacctgacagagaaacttaaccgagcagagagcggccacaacactcctaaagggaag	4213
QY	4018		gccaggtctgggaagacctgactggctttaagagctcttccagaccctctgggtcact	4077
Db	4214		gccaggtctgggaagacctgactggctttaagagctcttccagaccctctgggtcact	4273
QY	4078		gaagaagcagtggtgtgtgcaaaaacttaataatgccctgcgaatcttctccaccagaa	4137
Db	4274		gaagaagcagtggtgtgtgcaaaaacttaataatgccctgcgaatcttctccaccagaa	4333
QY	4138		tcagcagacaccccaacaagcacaagaagcgaccgccaagacaccttttgagaaaaagggac	4197
Db	4334		tcagcagacaccccaacaagcacaagaagcgaccgccaagacaccttttgagaaaaagggac	4393
QY	4198		gtacagaaaggagctctcagccctgaagaggtcacacagacatcaggggaaaaaccacac	4257
Db	4394		gtacagaaaggagctctcagccctgaagaggtcacacagacatcaggggaaaaaccacac	4453
QY	4258		acagataaagatccagaggtgaggaataaagcatcaacgcgttttagggaaactgcacaa	4317
Db	4454		acagataaagatccagaggtgaggaataaagcatcaacgcgttttagggaaactgcacaa	4513

QY 4318 cagaactggacccagcagcaagtgttaactggttagcaagaggcaccacaaactaagaaa 4377
DB 4514 cagaaactggacccagcagcaagtgttaactggttagcaagaggcaccacaaactaagaaa 4573
QY 4378 aaggcccaacccctagaagaactggctggctggaagaagctcttccagacaccagtatgc 4437
DB 4574 aaggcccaacccctagaagaactggctggctggaagaagctcttccagacaccagtatgc 4633
QY 4438 actgacaagcccaagactcacgagaaaactaccaaaatagcctgcagatcacaaaccagac 4497
DB 4634 actgacaagcccaagactcacgagaaaactaccaaaatagcctgcagatcacaaaccagac 4693
QY 4498 ccagtggaacacacaaacagctccaagccacagctccaagagaagctctcaggaaaagtggac 4557
DB 4694 ccagtggaacacacaaacagctccaagccacagctccaagagaagctctcaggaaaagtggac 4753
QY 4558 gtagaagaagaattcttcgcactcaggaaaacgaacacacatcagcaggcaagccaatgcac 4617
DB 4754 gtagaagaagaattcttcgcactcaggaaaacgaacacacatcagcaggcaagccaatgcac 4813
QY 4618 acacccaaccagcagtaagtgggtgagaaaacatctacgcatttatgggaactccagtgc 4677
DB 4814 acacccaaccagcagtaagtgggtgagaaaacatctacgcatttatgggaactccagtgc 4873
QY 4678 cagaaactggacctgacagagaacttaactggcagcaagagacggctacaaactccttaag 4737
DB 4874 cagaaactggacctgacagagaacttaactggcagcaagagacggctacaaactccttaag 4933
QY 4738 gaaaaggcccggtctctagaagacctggctggctttaaagagctcttcacagacacagagt 4797
DB 4934 gaaaaggcccggtctctagaagacctggctggctttaaagagctcttcacagacacagagt 4993
QY 4798 cacactgagggaatcaatgactaaacgataaaaactccaagtagcctgcaaatcttcacaa 4857
DB 4994 cacactgagggaatcaatgactaaacgataaaaactccaagtagcctgcaaatcttcacaa 5053
QY 4858 ccagacctagacaaaaacccagcaagctccaagcgacggctcaagacatccctggggaaa 4917
DB 5054 ccagacctagacaaaaacccagcaagctccaagcgacggctcaagacatccctggggaaa 5113
QY 4918 gtggcggtgaagaagaactcctagcagttggcaagctccacagacatcaggagagact 4977
DB 5114 gtggcggtgaagaagaactcctagcagttggcaagctccacagacatcaggagagact 5173
QY 4978 acacacacacacagagcccaacagggagatggttaagagatgaaagcatttatggagtct 5037
DB 5174 acacacacacacagagcccaacagggagatggttaagagatgaaagcatttatggagtct 5233
QY 5038 ccaagcgagatcttagactcagcagcaagtcttaactggcagcaagaggcagctgagaact 5097
DB 5234 ccaagcgagatcttagactcagcagcaagtcttaactggcagcaagaggcagctgagaact 5293
QY 5098 cctaaaggaaagtctgaagtccctgaagaactggccggtctcatcagactcttcacagaca 5157
DB 5294 cctaaaggaaagtctgaagtccctgaagaactggccggtctcatcagactcttcacagaca 5353
QY 5158 ccaagtcacactaagggaatcaatgactaatgaaaaactaccaaagtattctctacagagct 5217
DB 5354 ccaagtcacactaagggaatcaatgactaatgaaaaactaccaaagtattctctacagagct 5413
QY 5218 tcacagccagactagtggacaccccaacaaagctccaagccacagcccaagagaagctc 5277
DB 5414 tcacagccagactagtggacaccccaacaaagctccaagccacagcccaagagaagctc 5473
QY 5278 aggaagcagacactgaagaagaatttttagcattttagaaacaaaacccaatcagcagc 5337
DB 5474 aggaagcagacactgaagaagaatttttagcatttttagaaacaaaacccaatcagcagc 5533
QY 5338 aaagccatgcacacaccccaacccagcagtaggtgaagagaagacatcaacacgtttttg 5397
DB 5534 aaagccatgcacacaccccaacccagcagtaggtgaagagaagacatcaacacgtttttg 5593

QY 5398 ggaactccagtcgcaaaaactggaccagccaggaataatttacctggcagcaataagacggcta 5457
DB 5594 ggaactccagtcgcaaaaactggaccagccaggaataatttacctggcagcaataagacggcta 5653
QY 5458 caaactcgttaaggaaaaagcccgagctctagaagaactgcactgcgttcagagagcttttc 5517
DB 5654 caaactcgttaaggaaaaagcccgagctctagaagaactgcactgcgttcagagagcttttc 5713
QY 5518 cagacacatgcactgtataacccccacagctgatgaaaaactaccaaaaaatactctgc 5577
DB 5714 cagacacatgcactgtataacccccacagctgatgaaaaactaccaaaaaatactctgc 5773
QY 5578 aaatccgcgaatcagaccccgagcgaccccccaacaaacacaaagcaacggcccaagaga 5637
DB 5774 aaatccgcgaatcagaccccgagcgaccccccaacaaacacaaagcaacggcccaagaga 5833
QY 5638 agcctcaagaagaagcagcgttagaggaagaatttttagcaattcaggaaaactaacacatca 5697
DB 5834 agcctcaagaagaagcagcgttagaggaagaatttttagcaattcaggaaaactaacacatca 5893
QY 5698 gcaggcaaaagcccatgcacacgcttaaaagcagcagtaggtgaagagaagaacatcaacaca 5757
DB 5894 gcaggcaaaagcccatgcacacgcttaaaagcagcagtaggtgaagagaagaacatcaacaca 5953
QY 5758 ttgtggggaactccagtggaactggaaactggacctgctaggaattttacctggcgagcaagaga 5817
DB 5954 ttgtggggaactccagtggaactggaaactggacctgctaggaattttacctggcgagcaagaga 6013
QY 5818 cggccacaaactccttaagaaaaagcccaaggtccttagaagatctcgtggttctaaaagag 5877
DB 6014 cggccacaaactccttaagaaaaagcccaaggtccttagaagatctcgtggttctaaaagag 6073
QY 5878 ctctccagacacaccaggttcacactgagaaatcaatgaccgagtcaaaaatcacaaagta 5937
DB 6074 ctctccagacacaccaggttcacactgagaaatcaatgaccgagtcaaaaatcacaaagta 6133
QY 5938 tctgcaaatctcccaacacccagaccagtccaaaaccccaacaaagctccaagcaacgactc 5997
DB 6134 tctgcaaatctcccaacacccagaccagtccaaaaccccaacaaagctccaagcaacgactc 6193
QY 5998 aagatatccttggggaagaagttagtgtgaagaagagggtcctaccagtcgggaagctcaca 6057
DB 6194 aagatatccttggggaagaagttagtgtgaagaagagggtcctaccagtcgggaagctcaca 6253
QY 6058 cagacgtcagggaagaacacacacacagacagacagcagcagcagcagcagcagcagcagc 6117
DB 6254 cagacgtcagggaagaacacacacacagacacacagacagacagcagcagcagcagcagcagc 6313
QY 6118 aaagcgtttaagggaatctgcgaagcagatgctggaccacagcaaaactatggaactgggagt 6177
DB 6314 aaagcgtttaagggaatctgcgaagcagatgctggaccacagcaaaactatggaactgggagt 6373
QY 6178 gagaagtgcccaagaacacacctaaaggaagggcccaatcactagaagacactggccggcttc 6237
DB 6374 gagaagtgcccaagaacacacctaaaggaagggcccaatcactagaagacactggccggcttc 6433
QY 6238 aaagagctcttcagacacccagacacacactgagggaatcaacaactgatacaaaaaactacc 6297
DB 6434 aaagagctcttcagacacccagacacacactgagggaatcaacaactgatacaaaaaactacc 6493
QY 6298 aaaa tagcctgcgaatctccaccacacagaatcaatggacactcccaacagcacaagagg 6357
DB 6494 aaaa tagcctgcgaatctccaccacacagaatcaatggacactcccaacagcacaagagg 6553
QY 6358 cggcccaaaaacaccttttggggaagggatatagttgggaagagctctcagccctgaagcag 6417
DB 6554 cggcccaaaaacaccttttggggaagggatatagttgggaagagctctcagccctgaagcag 6613
QY 6418 ctcacacagacacacacacagacaaagtaccagagataagagataaaagcgcacacgtg 6477
DB 6614 ctcacagacacacacacacagacaaagtaccagagataagagataaaagcgcacacgtg 6573
QY 6478 ttcaagggaactgcaaaaacacagaaactggaccacagcaagtgtaactggttagcagaagg 6537

Db	6674	ttcagggaactgcataaacgaactggaccagcagcaagttaactggtagcaagg	6733
QY	6538	cagccagaactccttaagggaagcccaacccttagaagacttggctggttgaagag	6597
Db	6734	cagccaagaactccttaagggaagcccaacccttagaagacttggctggttgaagag	6793
QY	6598	ctcttcagacacacagtatgactgacaagcccacgactcacgagaaaactaccaata	6657
Db	6794	ctcttcagacacacagtatgactgacaagcccacgactcacgagaaaactaccaata	6853
QY	6658	gctcgagatctccaaaccagaccagtggttaccccaacaattcttcaaggccacagtc	6717
Db	6854	gctcgagatctccaaaccagaccagtggttaccccaacaattcttcaaggccacagtc	6913
QY	6718	aagagaagtcctcaggaagcagcgttagaggaagaatccttagcactcaggaaacga	6777
Db	6914	aagagaagtcctcaggaagcagcgttagaggaagaatccttagcactcaggaaacga	6973
QY	6778	ccatcagtagggaagcctaaggacacaccccaaccagcggaggtgatgagaagacatg	6837
Db	6974	ccatcagtagggaagcctaaggacacaccccaaccagcggaggtgatgagaagacatg	7033
QY	6838	aaagcatttatgggaactccagtcgagaaattggacctgccaggaattttacctggcgc	6897
Db	7034	aaagcatttatgggaactccagtcgagaaattggacctgccaggaattttacctggcgc	7093
QY	6898	aaagaatggccacaactcctaaggaaagcccgagctctagaagacctggctgcttc	6957
Db	7094	aaagaatggccacaactcctaaggaaagcccgagctctagaagacctggctgcttc	7153
QY	6958	aaagagctcttcagacacacagcactgacaagcccacgactgatgaaaaactacaaa	7017
Db	7154	aaagagctcttcagacacacagcactgacaagcccacgactgatgaaaaactacaaa	7213
QY	7018	atagctgcaaatctccaaaccagaccagtggaacaccccaagcagacaaagcaacgg	7077
Db	7214	atagctgcaaatctccaaaccagaccagtggaacaccccaagcagacaaagcaacgg	7273
QY	7078	cccaagagaactcctcaggaagcagcgttagaggaagaatttttagcactcaggaaacga	7137
Db	7274	cccaagagaactcctcaggaagcagcgttagaggaagaatttttagcactcaggaaacga	7333
QY	7138	acaccatcagcaggaagccatggaaccccccaaccagcagtaagttagagaaaaat	7197
Db	7334	acaccatcagcaggaagccatggaaccccccaaccagcagtaagttagagaaaaat	7393
QY	7198	atcaacacatttgtgaaactccagtcgagaaactggacctgtaggaaattttacctggc	7257
Db	7394	atcaacacatttgtgaaactccagtcgagaaactggacctgtaggaaattttacctggc	7453
QY	7258	agcaagagacagccacagactccttaaggaaagcctgagccttagagacctgggtggc	7317
Db	7454	agcaagagacagccacagactccttaaggaaagcctgagccttagagacctgggtggc	7513
QY	7318	ttcaagaactcttcagacacacagctcacactgaggaataatgactgatgacaaaatc	7377
Db	7514	ttcaagaactcttcagacacacagctcacactgaggaataatgactgatgacaaaatc	7573
QY	7378	acagaagtatctgtaaatctccacagccagagtcatttcaaaacctcaagaagctccaag	7437
Db	7574	acagaagtatctgtaaatctccacagccagagtcatttcaaaacctcaagaagctccaag	7633
QY	7438	caaaagctcaagatacccttggtaagtggacatgaaagaagagcccttagcagtcagc	7497
Db	7634	caaaagctcaagatacccttggtaagtggacatgaaagaagagcccttagcagtcagc	7693
QY	7498	aagctcacacagacatacaggggagactacgcgcaaacacacacagagccaacagagatag	7557
Db	7694	aagctcacacagacatacaggggagactacgcgcaaacacacacagagccaacagagatag	7753
QY	7558	aagagcatcaagcgttttaaggagctccaaagcagactcctggaccacagcaagtgta	7617
Db			

Db	7754	aagagcatcaaaagcgtttaaggagctctccaaagcagatccttggaaccagcagcaagtga	7813
QY	7618	actggtagcagcagcagcgtgagaactcgttaagaaaaagcccgctctctagaagacctg	7677
Db	7814	actggtagcagcagcagcgtgagaactcgttaagaaaaagcccgctctctagaagacctg	7873
QY	7678	gttgacttcaaaagcgtctctcagcaccaggttcacactgaagagtcgaatgactattgac	7737
Db	7874	gttgacttcaaaagcgtctctcagcaccaggttcacactgaagagtcgaatgactattgac	7933
QY	7738	aaaaaacacaaaaattccctgcataattctccccaccagaaactaacagacactgccacagc	7797
Db	7934	aaaaaacacaaaaattccctgcataattctccccaccagaaactaacagacactgccacagc	7993
QY	7798	acaaagagatgcccccaagacacgtlcccgagaaagagtaaaagagagctctcagcagtt	7857
Db	7994	acaaagagatgcccccaagacacgtlcccgagaaagagtaaaagagagctctcagcagtt	8053
QY	7858	gagagctcgcgcaaacatcagggcgaagcacacacacacacacacacacacacacacac	7917
Db	8054	gagagctcgcgcaaacatcagggcgaagcacacacacacacacacacacacacacacac	8113
QY	7918	gatgaagggcatcaaaagtattgaagcaacgtgcgaaagaagaaccccaaccagtagaagag	7977
Db	8114	gatgaagggcatcaaaagtattgaagcaacgtgcgaaagaagaaccccaaccagtagaagag	8173
QY	7978	gaaccagcagcaggaagcccaagcaccctaaagaaaaagcccaaccctgggaagacctg	8037
Db	8174	gaaccagcagcaggaagcccaagcaccctaaagaaaaagcccaaccctgggaagacctg	8233
QY	8038	gcgggtttcacagagctctctgaacatcaggttcacactcaggaatcactgactgctggc	8097
Db	8234	gcgggtttcacagagctctctgaacatcaggttcacactcaggaatcactgactgctggc	8293
QY	8098	aaagccactaaaaatccctgcgaaatctcccccactagaagtggtagacacacagcaagc	8157
Db	8294	aaagccactaaaaatccctgcgaaatctcccccactagaagtggtagacacacagcaagc	8353
QY	8158	acaaagggcatctcaggacacgtgtgcagaaggtacaagtataaagaagagccttcagca	8217
Db	8354	acaaagggcatctcaggacacgtgtgcagaaggtacaagtataaagaagagccttcagca	8413
QY	8218	gtcaagtccaacacatcaggggaaacccacggtatgcagacaaagaacccagcaggtgaa	8277
Db	8414	gtcaagtccaacacatcaggggaaacccacggtatgcagacaaagaacccagcaggtgaa	8473
QY	8278	gataaaggcatcaaaagcattgaaggaaatctgcgaaacagacacccggtctcagcagcaag	8337
Db	8474	gataaaggcatcaaaagcattgaaggaaatctgcgaaacagacacccggtctcagcagcaag	8393
QY	8338	gtaactggcagcagcagcagcccaagcagcaccagggaaagtccccaaagcca tagaagac	8397
Db	8534	gtaactggcagcagcagcagcccaagcagcaccagggaaagtccccaaagcca tagaagac	8593
QY	8398	ctagctggcttcaaaagacccagcagcaggttcacactgaagaaatcaatgactgtagacaaa	8457
Db	8594	ctagctggcttcaaaagacccagcagcaggttcacactgaagaaatcaatgactgtagacaaa	8653
QY	8458	accactaaaaatccctgcgaaatctcaccagaaactcaccagaaactcaccagaaactc	8517
Db	8654	accactaaaaatccctgcgaaatctcaccagaaactcaccagaaactcaccagaaactc	8713
QY	8518	agacggcccgagcagcagtcgccgaaagtagaagtgaaggagagctgttagcagttggc	8577
Db	8714	agacggcccgagcagcagtcgccgaaagtagaagtgaaggagagctgttagcagttggc	8773
QY	8578	aagctcacacaaacctcaggggagacacgcacacacacacacacacacacacacacacac	8637
Db	8774	aagctcacacaaacctcaggggagacacgcacacacacacacacacacacacacacacac	8833
QY	8638	aaagcagcaaaagcatttaagcaaacctgaaaagcagcgtggacacagaaagatgtaatt	8697
Db	8834	aaagcagcaaaagcatttaagcaaacctgaaaagcagcgtggacacagaaagatgtaatt	8893

QY 8698 ggcagcagagacagcccaagagcacctaaggaaaaaggccccaccctctggaacacctggcc 8757
|||||
Db 8894 ggcagcagagacagcccaagagcacctaaggaaaaaggccccaccctctggaacacctggcc 8953
QY 8758 agcttccaaagagctctctcaaacacacagccacactgaggaaactggcaaatggtctgct 8817
|||||
Db 8954 agcttccaaagagctctctcaaacacacagccacactgaggaaactggcaaatggtctgct 9013
QY 8818 gatagttttaacagcgctccaaaagcaaacacactgacagtggaaaacctctaaaaataacc 8877
|||||
Db 9014 gatagttttaacagcgctccaaaagcaaacacactgacagtggaaaacctctaaaaataacc 9073
QY 8878 agaagagtcttcggggcccttaaaagtagaacccgtgggagacgtggtgaagcaccagagac 8937
|||||
Db 9074 agaagagtcttcggggcccttaaaagtagaacccgtgggagacgtggtgaagcaccagagac 9133
QY 8938 cctgtaaaatacacaagcaaaaagcaaacacttccctgcccccaactgccccttaagaaggga 8997
|||||
Db 9134 cctgtaaaatacacaagcaaaaagcaaacacttccctgcccccaactgccccttaagaaggga 9193
QY 8998 ggtggcaaatggaagcgctcacgggaaccaagagcgctgcgtcatgccagcaccagag 9057
|||||
Db 9194 ggtggcaaatggaagcgctcacgggaaccaagagcgctgcgtcatgccagcaccagag 9253
QY 9058 gaaattgtggaggagctgccagccagcaagaagcagaggggttgcctccagggccaagagcc 9117
|||||
Db 9254 gaaattgtggaggagctgccagccagcaagaagcagaggggttgcctccagggccaagagcc 9313
QY 9118 aaatcatccgaacccgtggtcatcatgaagagaagtgttgaggaacttctgcataaagaatt 9177
|||||
Db 9314 aaatcatccgaacccgtggtcatcatgaagagaagtgttgaggaacttctgcataaagaatt 9373
QY 9178 gaacctcggaagagctgaacagcaacgacatgaaaccccaacaaagagagaaacacaaatta 9237
|||||
Db 9374 gaacctcggaagagctgaacagcaacgacatgaaaccccaacaaagagagaaacacaaatta 9433
QY 9238 caagactcgctccctgaaaaataagggaatatccctgcgtccagcgcgaagataagact 9297
|||||
Db 9434 caagactcgctccctgaaaaataagggaatatccctgcgtccagcgcgaagataagact 9493
QY 9298 gaggcagaacagcaataactgaggtcttatttagcagaaaaagaataaataaacaga 9357
Db 9494 gaggcagaacagcaataactgaggtcttatttagcagaaaaagaataaataaacaga 9553
QY 9358 aatgaaaaagagcccatgaagacctccccagagatggacattcagaatccagatgatga 9417
|||||
Db 9534 aatgaaaaagagcccatgaagacctccccagagatggacattcagaatccagatgatga 9613
QY 9418 gcccgaaaaccccatacctagagacaaagtctactgagaaacaaaggctgtgaggtctgct 9477
|||||
Db 9614 gcccgaaaaccccatacctagagacaaagtctactgagaaacaaaggctgtgaggtctgct 9673
QY 9478 agacagaatgagagctccagcctaaggtggcagaggagagcgaggggcagaagagtgcg 9537
Db 9674 agacagaatgagagctccagcctaaggtggcagaggagagcgaggggcagaagagtgcg 9733
QY 9538 aaggttctcatgcagaatcagaaagggaaggagagcaggaattcagactcccatgtgc 9597
|||||
Db 9734 aaggttctcatgcagaatcagaaagggaaggagagcaggaattcagactcccatgtgc 9793
QY 9598 ctgagatcaagaaaaacaaaagccagctgcagcaaacactttggagagcaaatctgtg 9657
Db 9794 ctgagatcaagaaaaacaaaagccagctgcagcaaacactttggagagcaaatctgtg 9853
QY 9658 cagagagtaacgcgagtgctcaagggtgtgcagaaaaatccaaagaaggctgagacaat 9717
Db 9854 cagagagtaacgcgagtgctcaagggtgtgcagaaaaatccaaagaaggctgagacaat 9913
QY 9718 gtgtgtgtcaagaaataaacaccagaggtcatatgggacagtgaagata 9766
|||||
Db 9914 gtgtgtgtcaagaaataaacaccagaggtcatatgggacagtgaagata 9962

RESULT 3
AAA64659
ID AAA64659 standard; cDNA; 11435 BP.
XX
AC AAA64659;
XX
DT 02-JAN-2001 (first entry)
XX
DE cDNA sequence encoding the antigen of monoclonal antibody Ki-67.
XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW antibody Ki-67; ss.
XX
OS Homo sapiens.
XX
PN WO200050595-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODIN N.
PA (FILO/) FIILONENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
PI Gout J, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
XX
DR WPI; 2000-572092/53.
XX
PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
XX
PS Claim 17; Page 81-85; 94pp; English.
XX
CC The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 11435 BP; 3802 A; 2828 C; 2684 G; 2121 T; 0 other;

Query Match 84.8%; Score 8286; DB 21; Length 11435;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1481 atgagagtgaaggaataaccttgaaagagcggtgtgtcctttgtggtgacacctaagac 1540
|||||
Db 597 atgagagtgaaggaataaccttgaaagagcggtgtgtcctttgtggtgacacctaagac 656
QY 1541 ctgaactatttgataaaacttgcctcttaacgctctcctcaaaaggagagagcccca 1600
|||||
Db 657 ctgaactatttgataaaacttgcctcttaacgctctcctcaaaaggagagagcccca 716
QY 1601 ccaaaagaagtctctgttaatgcacactccacctctctctgaagaaaaatcataaggaac 1660
|||||
Db 717 ccaaaagaagtctctgttaatgcacactccacctctctctgaagaaaaatcataaggaac 776
QY 1661 agccctcaaccatcaggaaaaacaaagactcaggttcagaaatccatgtggaagtgaaggaac 1720
|||||
Db 777 agccctcaaccatcaggaaaaacaaagactcaggttcagaaatccatgtggaagtgaaggaac 836

QY 1721 aaagcttggtataagccctccagctcctagtcctagtgaaataactccagttgcccagtgatc 1780
Db 837 aaagcttggtataagccctccagctcctagtcctagtgaaataactccagttgcccagtgatc 896
QY 1781 aagccgttagttcctgcataaacagccctgcttccagcagcaaatctcacacagaggttc 1840
Db 897 aagccgttagttcctgcataaacagccctgcttccagcagcaaatctcacacagaggttc 956
QY 1841 ctaagagagagagagaaagtgtggcaacctgcttcaaaagagagtgtctatcagccgaa 1900
Db 957 ctaagagagagagagaaagtgtggcaacctgcttcaaaagagagtgtctatcagccgaa 1016
QY 1901 gtcaacatgatatatttacagatgatatgttccaaagagaaagtgtgtcttcggaagcaa 1960
Db 1017 gtcaacatgatatatttacagatgatatgttccaaagagaaagtgtgtcttcggaagcaa 1076
QY 1961 atctgattgttgcataaatcatggcgagatgataaaacttggtcataaacacacaaa 2020
Db 1077 atctgattgttgcataaatcatggcgagatgataaaacttggtcataaacacacaaa 1136
QY 2021 ctaaaagtcaataaacatgtgtcccaaggttcaatgaacaaaaggcacaagagacctgtcta 2080
Db 1137 ctaaaagtcaataaacatgtgtcccaaggttcaatgaacaaaaggcacaagagacctgtcta 1196
QY 2081 ctccaaagaagccgtgtggcgaggttcaacagttcaacttagtacaggccacgcaaatcttc 2140
Db 1197 ctccaaagaagccgtgtggcgaggttcaacagttcaacttagtacaggccacgcaaatcttc 1256
QY 2141 ctgtaccataataataggaagagctcatactgaaaagtacatgtcctgctcgacct 2200
Db 1257 ctgtaccataataataggaagagctcatactgaaaagcacatgtcctgctcgacct 1316
QY 2201 acagagtgctcaacaacttcatttccaaacaaaatggactttaaggaagatctttcag 2260
Db 1317 acagagtgctcaacaacttcatttccaaacaaaatggactttaaggaagatctttcag 1376
QY 2261 gaatagctgaaatgttcaagaccccgctgaagagcaacccgagttgacaagcacatgtc 2320
Db 1377 gaatagctgaaatgttcaagaccccgctgaagagcaacccgagttgacaagcacatgtc 1436
QY 2321 acatcgctatttcaaatctcagagaatttgccttggaacacagtttcaaggaactgtattcag 2380
Db 1437 acatcgctatttcaaatctcagagaatttgccttggaacacagtttcaaggaactgtattcag 1496
QY 2381 gagaaacacctgctcccccacctcagagagtttggaggaataatgtgttcttcagtgcac 2440
Db 1497 gagaaacacctgctcccccacctcagagagtttggaggaataatgtgttcttcagtgcac 1556
QY 2441 agaatgcagcaaaaacagccatctgataaatgtctctgcaagccctcccttaagacgagct 2500
Db 1557 agaatgcagcaaaaacagccatctgataaatgtctctgcaagccctcccttaagacgagct 1616
QY 2501 gtattagagaaaatgaaacgtagcaaaaacgcccaggaacacactacaaaatgactcttc 2560
Db 1617 gtattagagaaaatgaaacgtagcaaaaacgcccaggaacacactacaaaatgactcttc 1676
QY 2561 tggagacaaaacttcagatacttgacagagccttcaaaaacagtatccactgttaaca 2620
Db 1677 tggagacaaaacttcagatacttgacagagccttcaaaaacagtatccactgttaaca 1736
QY 2621 ggtcagaagaggtctacagagttcaggaataatacagaaagctacactgtggaaagtcaaggtg 2680
Db 1737 ggtcagaagaggtctacagagttcaggaataatacagaaagctacactgtggaaagtcaaggtg 1796
QY 2681 aagaacaaatacagaaaattgttgagtgcatacctctaaaagagaggtcagaagggcaacactac 2740
Db 1797 aagaacaaatacagaaaattgttgagtgcatacctctaaaagagaggtcagaagggcaacactac 1856
QY 2741 tacaacaaagagagagagagatgaaggaaaatagaagaaccttttgacataataagg 2800
Db 1857 tacaacaaagagagagagagatgaaggaaaatagaagaaccttttgacataataagg 1916
QY 2801 aaaaattgaattaaaagaaaacgataaaaagatgaaagcaatgaaagatcaagaagactt 2860

Db 1917 aaaaattgaattaaaagaaaacgataaaaagatgaaagcaatgaaagatcaagaactt 1976
QY 2861 gggggcagaaaatgtgcaccaatgtctgacctgacagacctcaagagcttccctgatacag 2920
Db 1977 gggggcagaaaatgtgcaccaatgtctgacctgacagacctcaagagcttccctgatacag 2036
QY 2921 aactcatgaaagacacagcgccgtggccagaatctctccaaacccaagatacatgcccaagg 2980
Db 2037 aactcatgaaagacacagcgccgtggccagaatctctccaaacccaagatacatgcccaagg 2096
QY 2981 caccaaagagtggaaaaggcaaaataactaaaaatgccttgcctgcccagtcattacaaccagaac 3040
Db 2097 caccaaagagtggaaaaggcaaaataactaaaaatgccttgcctgcccagtcattacaaccagaac 2156
QY 3041 caataaacaccccaacacacacacaaaacacagttgaaagcatcctctggggaagtaggtg 3100
Db 2157 caataaacaccccaacacacacacaaaacacagttgaaagcatcctctggggaagtaggtg 2216
QY 3101 tgaagaagagctccttagcagtcggcaagtccacagcgagctcagggagaccacgcaca 3160
Db 2217 tgaagaagagctccttagcagtcggcaagtccacagcgagctcagggagaccacgcaca 2276
QY 3161 cgacagagagccagcagggagatggcaagagcatcagaacgctttaagagagcttcccaaac 3220
Db 2277 cgacagagagccagcagggagatggcaagagcatcagaacgctttaagagagcttcccaaac 2336
QY 3221 agatccttgaccacagcagcccggtgaactggaatgaaagagtgcccaagacgccttaagg 3280
Db 2337 agatccttgaccacagcagcccggtgaactggaatgaaagagtgcccaagacgccttaagg 2396
QY 3281 aagagcccgactcactagaagacctggctgcttcaaaagagcttccagacaccaggtc 3340
Db 2397 aagagcccgactcactagaagacctggctgcttcaaaagagcttccagacaccaggtc 2456
QY 3341 cctctgaggaatacaatgactgatgagaaaactaccaaaaatagcctgcaaatctccaccac 3400
Db 2457 cctctgaggaatacaatgactgatgagaaaactaccaaaaatagcctgcaaatctccaccac 2516
QY 3401 cagaaatcagtgagacactccaaacagcaacaagcaatggccttaagagagcttccaggaag 3460
Db 2517 cagaaatcagtgagacactccaaacagcaacaagcaatggccttaagagagagcttccaggaag 2576
QY 3461 cagatgtagagggaagattcttagcactcaggaactaacaccatcagcaggggaaagcca 3520
Db 2577 cagatgtagagggaagattcttagcactcaggaactaacaccatcagcaggggaaagcca 2636
QY 3521 tgccttacgcccacacagcagaggttgatgagaaaagacattaaagcatttatgggaactc 3580
Db 2637 tgccttacgcccacacagcagaggttgatgagaaaagacattaaagcatttatgggaactc 2696
QY 3581 cagtcagaaaactggacctggcgaggaactttacctggcagcaaaaagacagctcacagactc 3640
Db 2697 cagtcagaaaactggacctggcgaggaactttacctggcagcaaaaagacagctcacagactc 2756
QY 3641 ctaaggaaaaggcccaggctctagaagacctggctgctttaaagagcttccagactc 3700
Db 2757 ctaaggaaaaggcccaggctctagaagacctggctgctttaaagagcttccagactc 2816
QY 3701 ctggttcacacccgaggaattagtgctgtgtgtaaaaccactaaaaataccctcgactctc 3760
Db 2817 ctggttcacacccgaggaattagtgctgtgtgtaaaaccactaaaaataccctcgactctc 2876
QY 3761 cacagtcagaccccgatggacaccccaacagcaacaagcaacgcccacaaagagagatca 3820
Db 2877 cacagtcagaccccgatggacaccccaacagcaacaagcaacgcccacaaagagagatca 2936
QY 3821 ggaagcagatgtagagggagaactcttagcgtgcagggaatcttaatgccatcagcaggca 3880
Db 2937 ggaagcagatgtagagggagaactcttagcgtgcagggaatcttaatgccatcagcaggca 2996
QY 3881 aagccatgcacacgacctaaacctcagtgaggtgaaagaaagacatcatcatattgttg 3940

Db 2997 aagccatgcacacgacctaaaccatcagtaggtgaaagaaagacatcatctatttggg 3056
Qy 3941 gaactccagtgacaaaactggacctgacagagaacttaacggcgacgaagagcgccac 4000
Db 3057 gaactccagtgacaaaactggacctgacagagaacttaacggcgacgaagagcgccac 3116
Qy 4001 aaactcctaaggaagagccagggctctggaagacctgactggctttaagagctcttcc 4060
Db 3117 aaactcctaaggaagagccagggctctggaagacctgactggctttaagagctcttcc 3176
Qy 4061 agacccttggtcatactgaagaagcagtggtgctggccaaactactataaatgacctgg 4120
Db 3177 agacccttggtcatactgaagaagcagtggtgctggccaaactactataaatgacctgg 3236
Qy 4121 aatcttctccaccagaaatcagcagacaccccaacaagcacagaagcgagcccaagacac 4180
Db 3237 aatcttctccaccagaaatcagcagacaccccaacaagcacagaagcgagcccaagacac 3296
Qy 4181 ctttggagaaaaggagacctacagaaggagctctcagccctgaagaagctcacacagacat 4240
Db 3297 ctttggagaaaaggagacctacagaaggagctctcagccctgaagaagctcacacagacat 3356
Qy 4241 caggggaaccacacacacagataaaagtaccagaggtgaggataaaagcatcaacgcgt 4300
Db 3357 caggggaaccacacacacagataaaagtaccagaggtgaggataaaagcatcaacgcgt 3416
Qy 4301 ttagggaactgcaaaacagaaaactggaccagcagcaagtgttaactggtagcaagagcg 4360
Db 3417 ttagggaactgcaaaacagaaaactggaccagcagcaagtgttaactggtagcaagagcg 3476
Qy 4361 acccaaaaacttaagaaaaggcccaaccccttagaagacctggctggctggaagagctct 4420
Db 3477 acccaaaaacttaagaaaaggcccaaccccttagaagacctggctggctggaagagctct 3536
Qy 4421 tccagacaccagtgactacagaagccacagactccagagaaactacaaaatagacct 4480
Db 3537 tccagacaccagtgactacagaagccacagactccagagaaactacaaaatagacct 3596
Qy 4481 gcagatcaaacaccagaccagtgacacaccaaagaactccaagccacagctccaagaaa 4540
Db 3597 gcagatcaaacaccagaccagtgacacaccaaagaactccaagccacagctccaagaaa 3656
Qy 4541 gtctcagaaaagtgagactagaagaagaattcttcgcactcaggaagaacacacacatcag 4600
Db 3657 gtctcagaaaagtgagactagaagaagaattcttcgcactcaggaagaacacacacatcag 3716
Qy 4601 caggcaaaagccatgcacacaccccaaacacagcagtaagtgtgagaaaacatctacgcgt 4660
Db 3717 caggcaaaagccatgcacacaccccaaacacagcagtaagtgtgagaaaacatctacgcgt 3776
Qy 4661 ttatgggaactccagtcagaaaactggacctgacagagaacttaactggcagcaagagac 4720
Db 3777 ttatgggaactccagtcagaaaactggacctgacagagaacttaactggcagcaagagac 3836
Qy 4721 ggctacaaaactcctaagaaaaggcccggtctctagaagacctggctggctttaaaagagc 4780
Db 3837 ggctacaaaactcctaagaaaaggcccggtctctagaagacctggctggctttaaaagagc 3896
Qy 4781 tcttcagacacagaggtcacactaggaatcaatgactaaacgataaaaactcccaagtag 4840
Db 3897 tcttcagacacagaggtcacactaggaatcaatgactaaacgataaaaactcccaagtag 3956
Qy 4841 cctgcaaatcttcaaacaccagaccttagacaaaaccccgcaagctcccaagcgagcggtca 4900
Db 3957 cctgcaaatcttcaaacaccagaccttagacaaaaccccgcaagctcccaagcgagcggtca 4016
Qy 4901 agacatccctggggaagtgggcggtgaagaagagctccttagcaggttggcagagctcacac 4960
Db 4017 agacatccctggggaagtgggcggtgaagaagagctccttagcaggttggcagagctcacac 4076
Qy 4961 agacatcaggagagactcacacacacacacagagcccaacagggagatggtgaagcatga 5020
Db 4077 agacatcaggagagactcacacacacacacagagcccaacagggagatggtgaagcatga 4136

Qy 5021 aagcatttatggagtctccaaagcagatcttagactcagcagaactctaaactggcagca 5080
Db 4137 aagcatttatggagtctccaaagcagatcttagactcagcagaactctaaactggcagca 4196
Qy 5081 agaggcagctgagaactccttaaggaaaagctgaaagtcctgaagtcctgaagacctggcggttca 5140
Db 4197 agaggcagctgagaactccttaaggaaaagctgaaagtcctgaagacctggcggttca 4256
Qy 5141 tggagcttccagacaccaaactcaactaaagaaatcaatgactaaatgaaaaactacca 5200
Db 4257 tggagcttccagacaccaaactcaactaaagaaatcaatgactaaatgaaaaactacca 4316
Qy 5201 aagtatcctacagagcttccacgccagacctagtgacaccccaacaagctccaagccac 5260
Db 4317 aagtatcctacagagcttccacgccagacctagtgacaccccaacaagctccaagccac 4376
Qy 5261 agcccaagagaagctctcaggaaaagcagacactgaaagaagaatttttagcatttaggaac 5320
Db 4377 agcccaagagaagctctcaggaaaagcagacactgaaagaagaatttttagcatttaggaac 4436
Qy 5321 aaacgccatcagcagcaaaagccatgcacacaccccaaccagcagtagtgaaagagaag 5380
Db 4437 aaacgccatcagcagcaaaagccatgcacacaccccaaccagcagtagtgaaagagaag 4496
Qy 5381 acatacaacagctttttgggaactccagtcgagaaaactggaccagccaggaattttacctg 5440
Db 4497 acatacaacagctttttgggaactccagtcgagaaaactggaccagccaggaattttacctg 4556
Qy 5441 gcagcaatagacgggtcaaaactcgtaaaggaaaagcccaggtctctagaaagactgactg 5500
Db 4557 gcagcaatagacgggtcaaaactcgtaaaggaaaagcccaggtctctagaaagactgactg 4616
Qy 5501 gcttcagagagctttccagacacccatgcactgataaacccacagctgatgaaaaacta 5560
Db 4617 gcttcagagagctttccagacacccatgcactgataaacccacagctgatgaaaaacta 4676
Qy 5561 ccaaaaaatactctgcataatctccgcaatcagaccccgagcagcccaacacacaaa 5620
Db 4677 ccaaaaaatactctgcataatctccgcaatcagaccccgagcagcccaacacacaaa 4736
Qy 5621 agcaacggcccaagagaagcctcaagaaaagcagtagagaagaagaatttttagcattca 5680
Db 4737 agcaacggcccaagagaagcctcaagaaaagcagtagagaagaagaatttttagcattca 4796
Qy 5681 ggaactaacacatcagcagcgaagccatgcacgcctaaagcagcagtaggtgaag 5740
Db 4797 ggaactaacacatcagcagcgaagccatgcacgcctaaagcagcagtaggtgaag 4856
Qy 5741 agaaagacatcaacacatttggggaactccagtgagaaaactggacctgtaggaatt 5800
Db 4857 agaaagacatcaacacatttggggaactccagtgagaaaactggacctgtaggaatt 4916
Qy 5801 tacttggaagaagaagcggccacaaaactccttaagaaaaagccaggctctagaagatc 5860
Db 4917 tacttggaagaagaagcggccacaaaactccttaagaaaaagccaggctctagaagatc 4976
Qy 5861 tggctggcttcaaaagagctcttccagacacccaggtcacactgagaatcaatgaccgatg 5920
Db 4977 tggctggcttcaaaagagctcttccagacacccaggtcacactgagaatcaatgaccgatg 5036
Qy 5921 acaaaatcacagaagatactctgcaaatctcccaacccagaccccgatcaaaaaccccaaaa 5980
Db 5037 acaaaatcacagaagatactctgcaaatctcccaacccagaccccgatcaaaaaccccaaaa 5096
Qy 5981 gctccaagcaacgactcaagatatctctggggaagtagtgtgaaagaaggggtcctac 6040
Db 5097 gctccaagcaacgactcaagatatctctggggaagtagtgtgaaagaaggggtcctac 5156
Qy 6041 cagtcggaagctcacagacgtcaggaagacacacacagacacagacagagagagcag 6100
Db 5157 cagtcggaagctcacagacgtcaggaagacacacacagacacagacagagagagcag 5216

QY	6101	gagatgaaagagc	atcaaaacgctttaaggaatctgcaaaagcagatgctggaacccagcaa	6160
Db	5217	gagatgaaagagc	atcaaaacgctttaaggaatctgcaaaagcagatgctggaacccagcaa	5276
QY	6161	actatggaacct	ggatgagagatggccaagaacaccttaaggaagggcccaatcactag	6220
Db	5277	actatggaacct	ggatgagagatggccaagaacaccttaaggaagggcccaatcactag	5336
QY	6221	aagacctggccg	gtctcaaaagagctctccagacacacagaccacctgaggaatacaaaa	6280
Db	5337	aagacctggccg	gtctcaaaagagctctccagacacacagaccacctgaggaatacaaaa	5396
QY	6281	ctgatgacaaaa	ctaccaaaatagctgcaaaatctccaccaccagaaatcaatggacactc	6340
Db	5397	ctgatgacaaaa	ctaccaaaatagctgcaaaatctccaccaccagaaatcaatggacactc	5456
QY	6341	caacaagcaaa	gagggcgcccaaaacacctttggggaagggatatagtgaagagc	6400
Db	5457	caacaagcaaa	gagggcgcccaaaacacctttggggaagggatatagtgaagagc	5516
QY	6401	tctcagccct	gaagcagctcacacagaccacacacagacaaagtaccagggagatgagg	6460
Db	5517	tctcagccct	gaagcagctcacacagaccacacacagacaaagtaccagggagatgagg	5576
QY	6461	ataaaggat	tcaacgctgttccagggaactgcataaaacagaaaactggaccacgacgactg	6520
Db	5577	ataaaggat	tcaacgctgttccagggaactgcataaaacagaaaactggaccacgacgactg	5636
QY	6521	taactggt	tagcaagagcagcgaagacctctaaagggaagcccaacccctagaagact	6580
Db	5637	taactggt	tagcaagagcagcgaagacctctaaagggaagcccaacccctagaagact	5696
QY	6581	tggctggct	tgaagagctcttcagacacacagatgacctgacaagcccaacgactcaag	6640
Db	5697	tggctggct	tgaagagctcttcagacacacagatgacctgacaagcccaacgactcaag	5756
QY	6641	agaaaact	accataaaatagcctgcagatctccacaacacagaccagctgggtaccaccaaaa	6700
Db	5757	agaaaact	accataaaatagcctgcagatctccacaacacagaccagctgggtaccaccaaaa	5816
QY	6701	tcttcaagc	acacagtcacagagagctctcaggaagcagacagctagaggaagaatccttag	6760
Db	5817	tcttcaagc	acacagtcacagagagctctcaggaagcagacagctagaggaagaatccttag	5876
QY	6761	cactcaggaa	cacacacatcagtagggaaagctatggacacaccccaacccagcaggag	6820
Db	5877	cactcaggaa	cacacacatcagtagggaaagctatggacacaccccaacccagcaggag	5936
QY	6821	gtgatgaga	aagacatgaaaagctttatgggaactccagtcgagaatgtggaacctgccag	6880
Db	5937	gtgatgaga	aagacatgaaaagctttatgggaactccagtcgagaatgtggaacctgccag	5996
QY	6881	gaatattac	ctggcagcaaaaagatggcccaaaactcctaaaggaaagggccagggctctag	6940
Db	5997	gaatattac	ctggcagcaaaaagatggcccaaaactcctaaaggaaagggccagggctctag	6056
QY	6941	aagacctgg	cttcaaaagagctctccagacacacagcagcactgacaagcccaagcactg	7000
Db	6057	aagacctgg	cttcaaaagagctctccagacacacagcagcactgacaagcccaagcactg	6116
QY	7001	atgagaaaa	actaccataaaatagcctgcaaaatctccacaacacagaccagtcggaaccccaag	7060
Db	6117	atgagaaaa	actaccataaaatagcctgcaaaatctccacaacacagaccagtcggaaccccaag	6176
QY	7061	caagcaaaa	gcaacggcccaagagaacacctcaggaagcagacgtagaggaagaatttt	7120
Db	6177	caagcaaaa	gcaacggcccaagagaacacctcaggaagcagacgtagaggaagaatttt	6236
QY	7121	tagcactcag	ggaacacacacatcagcaggcaaaagccatggacaccccaacccagcag	7180
Db	6237	tagcactcag	ggaacacacacatcagcaggcaaaagccatggacaccccaacccagcag	6296
QY	7181	taagtgat	gagaaaaataatacaacacatttgtggaaactccagtcgagaaactggacctgc	7240

Db	6297	taagtgat	gagaaaaataatacaacacatttgtggaaactccagtcgagaaactggacctgc	6356
QY	7241	taggaattt	aaccttgccagcaagagacagccacagactcctaaagaaaagcctgaggtcc	7300
Db	6357	taggaattt	aaccttgccagcaagagacagccacagactcctaaagaaaagcctgaggtcc	6416
QY	7301	tagagacct	gtgtgcttcaaaagaaactctccagacacccaggttcacactcgaggaatacaa	7360
Db	6417	tagagacct	gtgtgcttcaaaagaaactctccagacacccaggttcacactcgaggaatacaa	6476
QY	7361	tgactgat	gacaaaaatacaagaagtatcctgttaaaatctccacagccagatcattcaaaa	7420
Db	6477	tgactgat	gacaaaaatacaagaagtatcctgttaaaatctccacagccagatcattcaaaa	6536
QY	7421	cctcaga	aaagctccaagcaaaagctcaagataccccctggtgaaagtggacatgaagaag	7480
Db	6537	cctcaga	aaagctccaagcaaaagctcaagataccccctggtgaaagtggacatgaagaag	6596
QY	7481	agccctag	cagtcagcaagctcacacgggacatcaggggagactagcgaaacacacacag	7540
Db	6597	agccctag	cagtcagcaagctcacacgggacatcaggggagactagcgaaacacacacag	6656
QY	7541	agccaac	agagatagtaagagacataaaagcgtttaaggaggtctccaaaagcagatcctgg	7600
Db	6657	agccaac	agagatagtaagagacataaaagcgtttaaggaggtctccaaaagcagatcctgg	6716
QY	7601	accacag	caaatgttaactggttagcagagcagctgagaactcgttaagaaaagggccc	7660
Db	6717	accacag	caaatgttaactggttagcagagcagctgagaactcgttaagaaaagggccc	6776
QY	7661	gtgctcta	gaagacctgggtgacttcaaaagagctcttctcagcaccaggttcacacactgaag	7720
Db	6777	gtgctcta	gaagacctgggtgacttcaaaagagctcttctcagcaccaggttcacacactgaag	6836
QY	7721	agtcaat	tgactattgcataaaacacaaaaatccctgcataatccccccaccagaactaa	7780
Db	6837	agtcaat	tgactattgcataaaacacaaaaatccctgcataatccccccaccagaactaa	6896
QY	7781	cagacact	ccacagcaaaaagagatgccccagacagctcccaggaagagatgataaag	7840
Db	6897	cagacact	ccacagcaaaaagagatgccccagacagctcccaggaagagatgataaag	6956
QY	7841	aggagct	ctcagcagttgagaggtctacgcaaaacatcaggggcaaaagcacacacacaca	7900
Db	6957	aggagct	ctcagcagttgagaggtctacgcaaaacatcaggggcaaaagcacacacacaca	7016
QY	7901	aagaaac	cagcaagcgtgtatgagggcatcaaaagtattgaaagcaacgtgcaaaagaaaac	7960
Db	7017	aagaaac	cagcaagcgtgtatgagggcatcaaaagtattgaaagcaacgtgcaaaagaaaac	7076
QY	7961	caaaac	ccagtagaagaggaaacccagcaggaaggccaagagcaccttaagggaagggccc	8020
Db	7077	caaaac	ccagtagaagaggaaacccagcaggaaggccaagagcaccttaagggaagggccc	7136
QY	8021	aacccct	ggaagacctggccggtctcacagagctctctgaaacatcaggttcacacactcagg	8080
Db	7137	aacccct	ggaagacctggccggtctcacagagctctctgaaacatcaggttcacacactcagg	7196
QY	8081	aatcact	gactgctggcaaaagcccaactaaaaataacctgcgaaatctccccactaagaagtgg	8140
Db	7197	aatcact	gactgctggcaaaagcccaactaaaaataacctgcgaaatctccccactaagaagtgg	7256
QY	8141	tagacac	ccacagcaaaaagagcactctcagggacagctgctgcaagaaggttcacaagtaa	8200
Db	7257	tagacac	ccacagcaaaaagagcactctcagggacagctgctgcaagaaggttcacaagtaa	7316
QY	8201	aagaa	gagccttcagcagtcgaagtccacacaaacatcagggggaacccagcgatgcagaca	8260
Db	7317	aagaa	gagccttcagcagtcgaagtccacacaaacatcagggggaacccagcgatgcagaca	7376
QY	8261	aagaa	ccacagcaggtggaagataaaaggcatcaaaagcattggaaggaattctgcaaaacagacac	8320

Db 7377 aagaaccagcagggtgaagataaaggcatcaaaagcattgaaggaatctctcaaaaacagacac 7436
Qy 8321 cggctccagcagcaagttaactggcagcagagagacgcccgaagacaccccgaggaaagt 8380
Db 7437 cggctccagcagcaagttaactggcagcagagagacgcccgaagacaccccgaggaaagt 7496
Qy 8381 cccaagccatagaagacactagctggcttcaaaagaccccgagcaggtcacactgaagaat 8440
Db 7497 cccaagccatagaagacactagctggcttcaaaagaccccgagcaggtcacactgaagaat 7556
Qy 8441 caatgactgatgacaaaaaccataaataccctgcaaaatcatcaccagaactagaaagaca 8500
Db 7557 caatgactgatgacaaaaaccataaataccctgcaaaatcatcaccagaactagaaagaca 7616
Qy 8501 ccgcaacaagctcaaaagacgcccaggaacacgtgcccagaaagtagaagtgaaggagg 8560
Db 7617 ccgcaacaagctcaaaagacgcccaggaacacgtgcccagaaagtagaagtgaaggagg 7676
Qy 8561 agctgttagcagttggcgaagctcacacaaacctcaggggagaccacgcacaccgcacaag 8620
Db 7677 agctgttagcagttggcgaagctcacacaaacctcaggggagaccacgcacaccgcacaag 7736
Qy 8621 agccggtaggtgaggcgaagacgcaaaagcatttaagcacaacctgcaagcggaacgtg 8680
Db 7737 agccggtaggtgaggcgaagacgcaaaagcatttaagcacaacctgcaagcggaacgtg 7796
Qy 8681 acgcagaagatgtaattggcagcagagacagcgaacacacttaagcacaacctgcaagcggaacgtg 8740
Db 7797 acgcagaagatgtaattggcagcagagacagcgaacacacttaagcacaacctgcaagcggaacgtg 7856
Qy 8741 cccctggaagacctggccagcttccaagagctctccaagacccagggccacactcagggaac 8800
Db 7857 cccctggaagacctggccagcttccaagagctctccaagacccagggccacactcagggaac 7916
Qy 8801 tggcaaatgtgctgctgaacttacaagcgtcccaagcgtcccaagcacaacacactgaagtgaa 8860
Db 7917 tggcaaatgtgctgctgaacttacaagcgtcccaagcgtcccaagcacaacacactgaagtgaa 7976
Qy 8861 aacctctaaaaatccagaagagtcttcgggcccctaaagttagaacccgtgggagacg 8920
Db 7977 aacctctaaaaatccagaagagtcttcgggcccctaaagttagaacccgtgggagacg 8036
Qy 8921 tggtaagcaccagagacctgtaaaaatcacaaagcacaacacttccctgccccac 8980
Db 8037 tggtaagcaccagagacctgtaaaaatcacaaagcacaacacttccctgccccac 8096
Qy 8981 tgccttcaagaggggaggtggcaaaagtgaagcgtcacgggaaccagagaggtgcgct 9040
Db 8097 tgccttcaagaggggaggtggcaaaagtgaagcgtcacgggaaccagagaggtgcgct 8156
Qy 9041 gcatgccagcaccagagaaattgtggagagctgccagccagcaagacagaggggtg 9100
Db 8157 gcatgccagcaccagagaaattgtggagagctgccagccagcaagacagaggggtg 8216
Qy 9101 ctccagggcaagaggcaaatcatccgaacccctggtcatcatgaagagaagtttgagga 9160
Db 8217 ctccagggcaagaggcaaatcatccgaacccctggtcatcatgaagagaagtttgagga 8276
Qy 9161 ctctgcaaaaaagaattgaacctgcggaagactgaacagcaacagacatgaaaaacca 9220
Db 8277 ctctgcaaaaaagaattgaacctgcggaagactgaacagcaacagacatgaaaaacca 8336
Qy 9221 aagaggcaacaaattacaagactcgtccctgaaaaaaggggaatatccctggcctcca 9280
Db 8337 aagaggcaacaaattacaagactcgtccctgaaaaaaggggaatatccctggcctcca 8396
Qy 9281 gacgccaagataaagactgaggcagaacagcaaaataactagtgaggtcttttattagcagaaa 9340
Db 8397 gacgccaagataaagactgaggcagaacagcaaaataactagtgaggtcttttattagcagaaa 8456
Qy 9341 gaatagaataaacaagaataaagaagcccatgaagacctccccagagatggacattc 9400
Db 8457 gaatagaataaacaagaataaagaagcccatgaagacctccccagagatggacattc 8516

Qy 9401 agaattccagatgatgagcccgaaaaccatacctagagacaaagtctactgagaacaaaa 9460
Db 8517 agaattccagatgatgagcccgaaaaccatacctagagacaaagtctactgagaacaaaa 8576
Qy 9461 ggtgtttaggtctctctagacagaaatgagagctccagcctaaagtggcagaggagagcg 9520
Db 8577 ggtgtttaggtctctctagacagaaatgagagctccagcctaaagtggcagaggagagcg 8636
Qy 9521 gagggcagaagagtcggaaggttctcatgcaaatcagaagaagggaagagagcaggaa 9580
Db 8637 gagggcagaagagtcggaaggttctcatgcaaaatcagaagaagggaagagcaggaa 8696
Qy 9581 attcagactccatgtgcctgagatcaagaagacaaaaagccagcctgcagcaagcactt 9640
Db 8697 attcagactccatgtgcctgagatcaagaagacaaaaagccagcctgcagcaagcactt 8756
Qy 9641 tggagagcaaatctgtcagagagtaacgcggagtgtaagagggtgtgcagaaaaatccaa 9700
Db 8757 tggagagcaaatctgtcagagagtaacgcggagtgtaagagggtgtgcagaaaaatccaa 8816
Qy 9701 agaagctggaagacaatgtgtgtcaagaaaaataacaaccagaagtcatagggacagtg 9760
Db 8817 agaagctggaagacaatgtgtgtcaagaaaaataacaaccagaagtcatagggacagtg 8876
Qy 9761 aagata 9766
Db 8877 aagata 8882
RESULT 4
AAH20873
ID AAH20873 standard; DNA; 733 BP.
XX
AC AAH20873;
XX
DT 21-AUG-2001 (first entry)
XX
DE Human Ki-67 protein KON-21 DNA fragment.
XX
DE Ki-67 protein; human; KON-21; anticancer; anti-allergic; gene therapy;
KW immunomodulatory; anti-inflammatory; anti-rheumatic; cancer; allergy;
KW autoimmune disease; inflammation; rheumatic disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 60..725
FT /*tag= a
FT /product= "KON-21"
XX
PN WO200136629-A1.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-EP11482.
XX
PR 18-NOV-1999; 95DE-1055576.
XX
XX (BORS-) FORSCHUNGSZENTRUM BORSTEL.
PI Gerdes J, Scholzen T, Wohlenberg C;
XX
XX WPI; 2001-355634/37.
DR P-PSDB; AAB86191.
XX
PT System for transferring compounds between cells, useful e.g. for
PT delivery of drugs or gene therapy sequences, based on the C-terminal
PT fragment of human protein Ki-67 -
XX
XX Disclosure; Fig 1; 3lpp; German.
CC This invention describes a novel use of the C-terminal fragment (A) of

CC Ki-67 protein (or its parts, fragments or homologs) for transfer of
CC attached compounds (B) into cells, or for uptake of (B) by, or release of
CC (B) from cells. The Ki-67 C-terminal fragment, designated KON-21, has
CC anticancer, anti-allergic, immunomodulatory, anti-inflammatory and
CC anti-rheumatic activity. (A), or transfer proteins (I) containing (A) and
CC at least one other component, are used as carriers for other active
CC ingredients in pharmaceutical compositions and to produce gene therapy
CC compositions, for treatment or prevention of, specifically, cancer,
CC allergy, autoimmune diseases, inflammation and rheumatic diseases. The
CC transfer vehicles are of human origin, avoiding risks associated with
CC using viral carriers. This sequence encodes the KON-21 protein fragment
CC used in the method of the invention.
XX
SQ Sequence 733 BP; 271 A; 144 C; 199 G; 119 T; 0 other;

Query Match 6.4%; Score 629; DB 22; Length 733;
Best Local Similarity 100.0%; Pred. No. 3.6e-284; Indels 0; Gaps 0;
Matches 629; Conservative 0; Mismatches 0;
QY 9108 ggcaagaggcaaatccatccgaaccctgggtcatcatgaagagaagtgttgaggactctgc 9167
Db 62 ggcaagaggcaaatccatccgaaccctgggtcatcatgaagagaagtgttgaggactctgc 121
QY 9168 aaaaagaattgaacctgcggaagagctgaacagcaacgacatgaaaacccaagaagaga 9227
Db 122 aaaaagaattgaacctgcggaagagctgaacagcaacgacatgaaaacccaagaagaga 181
QY 9228 acacaattacaagactcgtccctgaataaaggaaatccctgcgtccagacgca 9287
Db 182 acacaattacaagactcgtccctgaataaaggaaatccctgcgtccagacgca 241
QY 9288 agataagactgagcgagaacagcaataactgaggtcttctgtattagcagaaagaataga 9347
Db 242 agataagactgagcgagaacagcaataactgaggtcttctgtattagcagaaagaataga 301
QY 9348 aataacagaaaatgaaaagaagcccatgaagacctccccagagatggaattcagaatcc 9407
Db 302 aataacagaaaatgaaaagaagcccatgaagacctccccagagatggaattcagaatcc 361
QY 9408 agatgatgagcccgaaacccatacctagagacaagaagtcactgagacaacaaagtgctt 9467
Db 362 agatgatgagcccgaaacccatacctagagacaagaagtcactgagacaacaaagtgctt 421
QY 9468 gaggtctgtacagaagaatgagagctcccagcctaaggtggcagagagcgagggca 9527
Db 422 gaggtctgtacagaagaatgagagctcccagcctaaggtggcagagagcgagggca 481
QY 9528 gaagagtgcgaaggttctcatgcagaatcagaagaagggaagagagcggaattcaga 9587
Db 482 gaagagtgcgaaggttctcatgcagaatcagaagaagggaagagagcggaattcaga 541
QY 9588 ctccatgtcctgagatcaagaagaacaaaagccagcctgcagcaagcactttgagag 9647
Db 542 ctccatgtcctgagatcaagaagaacaaaagccagcctgcagcaagcactttgagag 601
QY 9648 caaatctgtcagagagtagtaacgcggaggtgtcaagaggtgtgcagaaaaatccaaagaagc 9707
Db 602 caaatctgtcagagagtagtaacgcggaggtgtcaagaggtgtgcagaaaaatccaaagaagc 661
QY 9708 tgaggacaatgtgtgttcaagaaaaataa 9736
Db 662 tgaggacaatgtgtgttcaagaaaaataa 690

RESULT 5
AAA64645
ID AAA64645 standard; DNA; 904 BP.
XX
AC AAA64645;
XX
DT 02-JAN-2001 (first entry)
XX

DE Partial sequence MEL10 of a Ki-67 nuclear antigen gene.
XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW Ki-67 nuclear antigen gene; ss.
XX
OS Homo sapiens.
XX
PN WO200050595-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000MO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODNIN N.
PA (FILO/) FILOENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
PI Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
XX
DR WPI; 2000-572092/53.
XX
PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
XX
PS Example 2; Page 66-67; 94pp; English.
XX
CC The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a partial sequence of a Ki-67 nuclear
CC antigen gene, and is a polynucleotide of the invention.
XX
SQ Sequence 904 BP; 318 A; 227 C; 221 G; 127 T; 11 other;

Query Match 5.4%; Score 527; DB 21; Length 904;
Best Local Similarity 99.7%; Pred. No. 2.3e-236;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6307 tgcaaatctccaccaccagaatcaatggacactccaacaagcacaaggaggcgcccaaa 6366
Db 276 tgcaaatctccaccaccagaatcaatggacactccaacaagcacaaggaggcgcccaaa 335
QY 6367 acaccttggggaaaaaggatagtaggaagagctctcagccctgaagcagctccacag 6426
Db 336 acaccttggggaaaaaggatagtaggaagagctctcagccctgaagcagctccacag 395
QY 6427 accacacacagacaaagtaccaggagatgaggaataaaggcacaacgtgttcaggga 6486
Db 396 accacacacagacaaagtaccaggagatgaggaataaaggcacaacgtgttcaggga 455
QY 6487 actcaaaaaacagaactggaccacagcagcaagtgttaactggtagcaaggaggcagcaaga 6546
Db 456 actcaaaaaacagaactggaccacagcagcaagtgttaactggtagcaaggaggcagcaaga 515
QY 6547 actcctaagggaaaaagcccaacccctagaaagacttgctggcttgaagaagctctccag 6606
Db 516 actcctaagggaaaaagcccaacccctagaaagacttgctggcttgaagaagctctccag 575
QY 6607 acaccagatgcacigacaaagcccagcactcacgagaaactaccacaaatagcctgcaga 6666
Db 576 acaccaataatgcactgacaaagcccagcactcacgagaaactaccacaaatagcctgcaga 635

QY 6667 tctccacaaccagaccagctgggtaccocccacaacaatcttccaaagccagtcctcaagaagaagt 6726
Db 636 tctccacaaccagaccagctgggtaccocccacaacaatcttccaaagccagtcctcaagaagaagt 695
QY 6727 ctccagaaagcagacgtagagaaagaatcctcttagcactcagcagaaacgaacacccatcagta 6786
Db 696 ctccagaaagcagacgtagagaaagaatcctcttagcactcagcagaaacgaacacccatcagta 755
QY 6787 gggaaagctatggacacaccccaaccagcagaggtgatgagaagaacagatgaaagcattt 6846
Db 756 gggaaagctatggacacaccccaaccagcagaggtgatgagaagaacagatgaaagcattt 815
QY 6847 atgggaactcagtcagaaattggacctgcccagagaaatttacctgggcagcaaaaagatgg 6906
Db 816 atgggaactcagtcagaaattggacctgcccagagaaatttacctgggcagcaaaaagatgg 875
QY 6907 ccacaactccttaagaaaagcccgagc 6935
Db 876 ccacaactccttaagaaaagcccgagc 904
RESULT 6
AAZ80374
ID AAZ80374 standard; cDNA; 561 BP.
XX
AC AAZ80374;
XX
DT 07-APR-2000 (first entry)
XX
DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:458.
XX
KW Human; gene expression product; diagnosis; tumour; colon cancer;
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW hyperplasia; ds.
XX
OS Homo sapiens.
XX
PN WO9964576-A2.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-IB01062.
XX
PR 10-JUN-1998; 98US-0088801.
XX
PA (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
PI Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
PI Schlegel R;
XX
DR WPI; 2000-087220/07.
XX
PT Novel nucleic acids, used to develop products for the diagnosis and
PT treatment of disorders involving unwanted cell proliferation,
PT particularly cancers, especially colon cancer -
PS Claim 15; Page 314; 469pp; English.
XX
CC AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g.
CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC can be used for developing agents for the diagnosis and treatment of
CC disorders involving unwanted cell proliferation, such as neoplasia,
CC dysplasia or hyperplasia.
XX
SQ Sequence 561 BP; 196 A; 143 C; 125 G; 81 T; 16 other;

Query Match 4.0%; Score 389; DB 21; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.1e-171;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6691 accccaacaatttcaagccacagtcctcaagagaagtctcaggaagacagacgtagaggaa 6750
Db 1 accccaacaatttcaagccacagtcctcaagagaagtctcaggaagacagacgtagaggaa 60
QY 6751 gaatcctagcactcaggaacgaacacacatcagtaggagaagctatgacacacccaaa 6810
Db 61 gaatccttagcactcaggaacgaacacacatcagtaggagaagctatgacacacccaaa 120
QY 6811 ccagcagagggtgatgagaagaacatgaaagcattttatggaaactccagtcagaaattg 6870
Db 121 ccagcagagggtgatgagaagaacatgaaagcattttatggaaactccagtcagaaattg 180
QY 6871 gacctgcaaggaatttacctggcagcaaaaagatggccaaactcctaaagaaaagccc 6930
Db 181 gacctgcaaggaatttacctggcagcaaaaagatggccaaactcctaaagaaaagccc 240
QY 6931 caggctctagaagacctggcttcaagagagctctccagacacccagcactgacaag 6990
Db 241 caggctctagaagacctggcttcaagagagctctccagacacccagcactgacaag 300
QY 6991 cccacgactgatgagaaaaactaccaaaatagcctgcaaatctccacaaccagccagtg 7050
Db 301 cccacgactgatgagaaaaactaccaaaatagcctgcaaatctccacaaccagccagtg 360
QY 7051 gacacccagcagcacaaaagcaacgccc 7079
Db 361 gacacccagcagcacaaaagcaacgccc 389
RESULT 7
AAA64643
ID AAA64643 standard; DNA; 850 BP.
XX
AC AAA64643;
XX
DT 02-JAN-2001 (first entry)
XX
DE Partial sequence MEL8 of a Ki-67 nuclear antigen gene.
XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW Ki-67 nuclear antigen gene; ss.
XX
OS Homo sapiens.
XX
PN WO200050595-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODIN N.
PA (FILO/) FILOENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
XX
DR WPI; 2000-572092/53.
XX
PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -

for use in e.g. forensics, paternity testing or for phenotypic typing of diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial hypercholesterolemia, polycystic kidney disease, hereditary spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, acute intermittent porphyria, autoimmune diseases, inflammation, cancer, diseases of the nervous system, infection by pathogenic microorganisms, and characteristics such as longevity, appearance (e.g. baldness, obesity), strength, speed, endurance, fertility, and susceptibility or receptivity to particular drugs or therapeutic treatments. The isolated polymorphic nucleic acid segments can also be used to produce medicaments for the treatment or prophylaxis of such diseases.

Sequence 201 BP; 35 A; 40 C; 48 G; 77 T; 1 Other;

Query Match 1.4%; Score 134; DB 19; Length 201;
Best Local Similarity 99.5%; Pred. No. 3.7e-52;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy 2430 cttcagtcgacagaatgcagcaaaacagccatctgtataatgtctgcgaagccctccctt 2489
Db 201 CTTTCAGTGCACAGAAATGCAGCAAAACAGCCATCTGTATAATGTCTGTCAAGCCCTCCCTT 142

Qy 2490 aagacgcacgtgtattagagaaaatggaacgtagcaaaacgcccaggaacacctacaa 2549
Db 141 AAGACGGCAGTGTATTAGAGAAAATGGAAACGTAGCAAAACGCCAGGAACACCTACAA 82

Qy 2550 aatgactctctggagacaaaacttcagatactgagacagagccttcaaaaacagctac 2609
Db 81 AATGACTTCTSTGGAGACAAAACTTCAGATACCTGAGACAGAGCCCTTCAAAAACAGTATC 22

Qy 2610 cactg 2614
Db 21 CACTG 17

RESULT 12
AAF67522
ID AAF67522 standard; cDNA; 158 BP.
XX AC AAF67522;
XX DT 09-APR-2001 (first entry)
XX Novel human polynucleotide, SEQ ID NO: 3284.
XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; ss.
XX Homo sapiens.
XX WO200102568-A2.
XX PN 2001-01-11
XX PD 11-JAN-2001.
XX PF 30-JUN-2000; 2000WO-0518374.
XX XX 02-JUL-1999; 99US-0142310.
XX PR 02-JUL-1999; 99US-0142311.
XX PA (CHIR) CHIRON CORP.
XX PA (HYSE-) HYSEQ INC.
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
XX Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
XX Ckenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
XX Kita D, Garcia V, Jones LW, Strache-Crain B;
XX WPI; 2001-091805/10.

Qy 8678 tggacgcagaagatgtaattggcagcagagacagccagagcaactaaggaaaggccc 8737
Db 266 tggacgcagaagatgtaattggcagcagagacagccagagcaactaaggaaaggccc 325

Qy 8738 aacccctggaagaccctggcagcttccaagagctctcacaacacagccacactgag 8797
Db 326 aacccctggaagaccctggcagcttccaagagctctcacaacacagccacactgag 385

Qy 8798 aactggcacaatggtgctgctgtagctttacaaagcgtccaaagcaaacacactgacagt 8857
Db 386 aactggcacaatggtgctgctgtagctttacaaagcgtccaaagcaaacacactgacagt 445

Qy 8858 gaaaacctctaaaatatccagaagagttcttctgggcccctaaagttagaacccgtggag 8917
Db 446 gaaaacctctaaaatatccagaagagttcttctgggcccctaaagttagaacccgtggag 505

Qy 8918 acgtggttaagcagacagacccctgttaaaatcacaagcaaaagcaaacacttccctgcc 8977
Db 506 acgtggttaagcagacagacccctgttaaaatcacaagcaaaagcaaacacttccctgcc 565

Qy 8978 cactgcccctcaagagggaggtgcaagatggaagcgtcacgggaacc 9027
Db 566 cactgcccctcaagagggaggtgcaagatggaagcgtcacgggaacc 615

RESULT 11
AAAX10717/c
ID AAAX10717 standard; DNA; 201 BP.
XX AC AAAX10717;
XX DT 30-MAR-1999 (first entry)
XX Human biallelic polymorphic DNA fragment WI-1011.
XX Polymorphism; biallelic; human; forensic; paternity testing; disease;
XX detection; phenotypic typing; characteristic; infection; hereditary;
XX autoimmune disease; cancer; inflammation; drug; therapy; medicament;
XX treatment; marker; ss.
XX Homo sapiens.
XX WO9820165-A2.
XX PD 14-MAY-1998.
XX PF 05-NOV-1997; 97WO-0520313.
XX XX 06-NOV-1996; 96US-0030455.
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX Hudson T, Lander ES, Wang D;
XX WPI; 1998-286974/25.
XX New isolated nucleic acid segments from the human genome - used for
XX determining polymorphic forms for use in e.g. forensics, paternity
XX testing or phenotypic typing for disease
XX Claim 1; Page 75; 310pp; English.
XX AAAX10269-X12937 are human DNA fragments which contain biallelic
XX polymorphic markers which have been isolated using the primers
XX represented in AAAX09121-X10268. The base occupying the polymorphic site
XX is indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
XX can be used in methods for determining polymorphic forms in an individual

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
PS Claim 9; Page 1036; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 158 BP; 54 A; 40 C; 34 G; 27 T; 3 other;

Query Match 1.1%; Score 110; DB 22; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.5e-41;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4699 aacttaactgcagcaagacgagcgtctacaactcctaaggaagccagcgtctagaa 4758
Db 11 aacttaactgcagcaagacgagcgtctacaactcctaaggaagccagcgtctagaa 70

Qy 4759 gacctggctgtttaaagagctctccagacagagtcacactgagga 4808
Db 71 gacctggctgtttaaagagctctccagacagagtcacactgagga 120

RESULT 13
AAT20857
ID AAT20857 standard; cDNA to mRNA; 219 BP.
XX
AC AAT20857;
XX
DT 10-JUL-1996 (first entry)
XX
DE Human gene signature HUMGS02107.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN WO9514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

XX
PS Claim 1; Page 743; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 219 BP; 93 A; 28 C; 57 G; 40 T; 1 other;

Query Match 0.8%; Score 78; DB 16; Length 219;
Best Local Similarity 100.0%; Pred. No. 6.5e-26;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9659 agagagtaacgcgcgagtggtcaagaggtgtgcagaaaaatccaaagagcgtgaggaatg 9718
Db 58 agagagtaacgcgcgagtggtcaagaggtgtgcagaaaaatccaaagagcgtgaggaatg 117

Qy 9719 tgtgtgtcaagaaaaataa 9736
Db 118 tgtgtgtcaagaaaaataa 135

RESULT 14
AAS69163
ID AAS69163 standard; cDNA; 1126 BP.
XX
AC AAS69163;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4967.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG04976.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 4967; 103pp; English.
XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1126 BP; 371 A; 274 C; 287 G; 194 T; 0 other;

Query Match 0.7%; Score 64; DB 23; Length 1126;
Best Local Similarity 100.0%; Pred. No. 2.5e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8752 ctggccagcttccaagagctctctcaaacaccagccacactgaggaaactggcgaatggt 8811

Db 379 ctggccagcttccaagagctctctcaaacaccagccacactgaggaaactggcgaatggt 438

QY 8812 gctg 8815

Db 439 gctg 442

RESULT 15

AAL27414

ID AAL27414 standard; DNA; 51 BP.

XX AAL27414;

XX 24-JAN-2002 (first entry)

XX Human SNP oligonucleotide #622.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; angiotensinogen; angiotensin converting enzyme;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.

XX Homo sapiens.

XX WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

XX 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX ShImkets RA, Leach M;

XX WPI; 2001-465210/50.

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
oncogenes and histones, useful for diagnosing and treating, e.g.
cancer, autoimmune diseases and infections -

Claim 1: Page 1565; 4143pp; English.

The present invention relates to oligonucleotides encoding polymorphic
variants of proteins related to amylases, amyloid proteins, angiotensin,
apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
histones, kinases, colony stimulating factors, complement related
proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
G-protein coupled receptors and thioesterases. The present sequence is
one such oligonucleotide. The oligonucleotides and the peptides encoded
by them may be used in the prevention, diagnosis and treatment of
diseases associated with inappropriate expression of the proteins listed
above. Disorders that may be prevented, diagnosed and/or treated include
multifactorial diseases with a genetic component, such as autoimmune
diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
systemic lupus erythematosus and Grave's disease), inflammation, cancer
(e.g. cancers of the bladder, brain, breast, colon and kidney,
leukaemia), diseases of the nervous system and an infection of pathogenic
organisms.

Sequence 51 BP; 20 A; 16 C; 8 G; 7 T; 0 other;

Query Match 0.5%; Score 51; DB 22; Length 51;

Best Local Similarity 100.0%; Pred. No. 2.9e-13;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4433 tatgcactgacagccacgactcacgagaaaactaccataatagctgcga 4483

Db 1 tatgcactgacagccacgactcacgagaaaactaccataatagctgcga 51

Search completed: August 24, 2002, 01:37:14

Job time: 28334 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 9568.17 Seconds
(without alignments)
13776.016 Million cell updates/sec

Title: US-09-700-906A-1_COPY_197_9962
Perfect score: 9766
Sequence: 1 atgtggccacgagacgct.....tcattaggagacagtaagata 9766

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	731	7.5	1058	10	BG389162 602414982
2	703	7.2	1084	10	BM455229 AGENCOURT
3	691	7.1	1022	10	BM474441 AGENCOURT
4	640	6.6	1009	10	BM459184 AGENCOURT
5	619	6.3	790	9	AL556595 AGENCOURT
6	576	5.9	1070	10	BM455791 AGENCOURT
7	575	5.9	697	10	BG403181 602418615
8	553	5.7	1068	10	BM452654 AGENCOURT
9	552	5.7	606	10	BE395002 601308294
10	544	5.6	580	9	AW835684 QV4-LT001
11	537	5.5	1047	10	BM476598 AGENCOURT
12	535	5.5	822	9	AU140843 AGENCOURT
13	525	5.4	753	10	BI868409 603392387
14	519	5.3	751	10	BG252689 602366335
15	510	5.2	692	10	BG677474 602625107
16	500	5.1	558	9	AW503684 UI-HF-BN0
17	497	5.1	546	9	AW502198 UI-HF-BR0

18	491	5.0	758	10	BG393561	BG393561 602411954
19	487	5.0	824	10	BE540030	BE540030 601060802
c 20	481	4.9	596	10	BF516576	BF516576 UI-H-BW1-
21	475	4.9	584	12	AQ471476	AQ471476 CITBI-E1-
22	471	4.8	471	10	BF898662	BF898662 QV1-MR022
23	459	4.7	596	9	AW502251	AW502251 UI-HF-BN0
24	442	4.5	852	10	BG613631	BG613631 602641986
c 25	439	4.5	599	9	A1064694	A1064694 HA0443 Hu
c 26	435	4.5	436	9	AA262266	AA262266 zs25e07.s
27	432	4.4	907	10	BG164890	BG164890 602343508
28	430	4.4	820	9	AU130074	AU130074 AU130074
29	421	4.3	505	9	AW504977	AW504977 UI-HF-BN0
30	421	4.3	683	9	AW853584	AW853584 RCL-CT025
c 31	420	4.3	557	9	AW835681	AW835681 QV4-LT001
c 32	416	4.3	446	9	A1494288	A1494288 QY98d08.x
33	414	4.2	561	9	AW499743	AW499743 UI-HF-BR0
34	414	4.2	667	10	BI861849	BI861849 603388877
35	413	4.2	639	9	BE083942	BE083942 PM0-BT065
36	411	4.2	672	9	AW853570	AW853570 RCL-CT025
37	406	4.2	471	9	AA262877	AA262877 zs25e07.r
38	403	4.1	851	10	BF966006	BF966006 602277193
c 39	401	4.1	465	9	AA809977	AA809977 oa95q07.s
40	397	4.1	960	10	BG032883	BG032883 602300466
41	396	4.1	508	9	A1002455	A1002455 oq82h11.s
42	394	4.0	688	10	BE740099	BE740099 601595030
c 43	388	4.0	912	9	AL562952	AL562952 AL562952
44	387	4.0	582	9	BE084144	BE084144 PM0-BT065
45	383	3.9	440	9	AW501940	AW501940 UI-HF-BR0

ALIGNMENTS

RESULT 1	BG389162	602414982F1	NIH_MGC_92	1058 bp	mRNA	linear	EST 12-MAR-2001
LOCUS	602414982F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4523108 5', mRNA sequence.						
DEFINITION	BG389162						
ACCESSION	BG389162.1						GI:13282608
VERSION	EST.						
KEYWORDS	human.						
SOURCE	Homo sapiens						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 1058)						
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.						
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)						
JOURNAL	Unpublished (1999)						
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov						

FEATURES

High quality sequence stop: 752.
Location/Qualifiers
1..1058
/db_xref="taxon:9606"
/clone="IMAGE:4523108"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT	404 a	258 c	252 g	144 t
ORIGIN				
Query Match	7.5%; Score 731; DB 10; Length 1058;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 731; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Qy 2653	cagaagctcacctgtggaagtgaagagtaagaacacaaatacagaaattgttgagtcactc 2712			
Db 1	CAGAAGCTCACCTGTGGAAAGTAAGAGTGAAGAACAAATACAGAATTTGTTCAGTGCATC 60			
Qy 2713	ctaaaaagagtgctgagaagggcaacactactacaacaaagagagaaggagatgaaggaa 2772			
Db 61	CTAAAAAGAGGTTCAGAAGGCAACACTACTACAACAAGGAGAGAAGGAGATGAAGGAA 120			
Qy 2773	atagaaagacacttttgagacataaagaaataattgaattaaaaagaaaaacgatgaagaag 2832			
Db 121	ATAGAAAGACCTTTTGAGACATATAAGGAAATATTGTAATTTAAAGAAACGATGAAAG 180			
Qy 2833	atgaagcaatgaagagatcaagaacttgggggcagaatgtgaccacaaatgtctgcacctg 2892			
Db 181	ATGAAGCAATGAGAGATCAAGAACTTGGGGCAGAAATGTGCACCAATGTCTGCACCTG 240			
Qy 2893	acagacctcaagagcttgcctgatacacagaactcatgaagacacagcgcgtggccagaat 2952			
Db 241	ACAGACCTCAAGAGCTTGCTTGATACAACTCATGAAAGACACGGCAGCTGGCCAGAAT 300			
Qy 2953	ctctccaaacccaaagatcatgcgaagggcaccaaagagtggaaggccaaatcactaaa 3012			
Db 301	CTCTCTCCAAACCCCAAGATCATGCCAAGGCACCAAGAGTGAGAAAGGCCAAATCACTAAA 360			
Qy 3013	atgcctctccagctcattacaaccagaaccataaacaccccaacacacacacacacacag 3072			
Db 361	ATGCCCTGCCAGTCATTACACACGAAACCATATAACACCCCAACACACACAAACACAG 420			
Qy 3073	ttgaaggcatccctgggggaagtgtgtgaaagaagagctcctagcagtcggcaagtcc 3132			
Db 421	TTGAAGGCATCCCTGGGGAAAGTAGTGTGAAAGAAGAGCTCCTAGCAGTCGGCAAGTTC 480			
Qy 3133	acaggaagctcaggggagaccacgcacacagagagaccagcagagagatggcaagagc 3192			
Db 481	ACAGCGAGCTCAGGGGAGACCCACCCACAGCAGAGAGCCAGCAGGAGATGGCAAGAGC 540			
Qy 3193	atcagaacgtttgaaggagctctccaaagcagatccttgaccacagcagccgtgttaactgga 3252			
Db 541	ATCAGACGTTTAAAGAGTCTCCAAAGCAGATCTCGACCCACGACGCCGCTGTAACTGGA 600			
Qy 3253	atgaagaagtggccaaagacccctaaagaaagagggcccgctcactagaagacctggctgac 3312			
Db 601	ATGAAGAAGTGGCCAAAGAACGCTTAAGGAAGAGGCCAGTCACTAGAAGACCTGGCTGGC 660			
Qy 3313	ttcaagaagctcttcagagacccaggtccctctcaggaatcaatgactgatgaaaaaact 3372			
Db 661	TTCAAAAGAGCTTCTCCAGACACACAGTCCCTCTGAGGAATCAATGACTGTGAGAAAACT 720			
Qy 3373	accaaaatagc 3383			
Db 721	ACCAAAATAGC 731			
RESULT 2				
BM455229				
LOCUS	BM455229 1084 bp mRNA linear EST 05-FEB-2002			
DEFINITION	AGENCOURT_6405486 NIH_MGC_85 Homo sapiens cdna clone IMAGE:5500179			
ACCSSION	5', mRNA sequence.			
VERSION	BM455229			
KEYWORDS	BM455229.1 GI:18504269			
SOURCE	EST.			
ORGANISM	human.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			

Db	541		CAGACATCAGAGAGACTACACACACACAGAGCCACAGAGATGGTAAAGCATG	600
Qy	5020	aaagcattatggagtgctcaaaagcagatcttagactcagcagcaagtctcaactggcagc	5079	
Db	601		AAAGCATTTATGGAGTCTCCAAAGCAGATCTTAGACTCAGCAGCAAGTCTAACTGGCAGC	660
Qy	5080	aagaggcagctgaaactcctaagggaaaagtctgaagctccctgaagacctggccggcttc	5139	
Db	661		AAGAGCGAGCTGAGAACTCCTAAGGGAAGTCTGAAGTCCTGAAGACCTGGCGGCTTC	720
Qy	5140	atgagctcttccagacacacaagtccacactaagg	5173	
Db	721		ATCAGAGCTCTCCAGACACCAAGTCACTAAGG	754
RESULT 3				
LOCUS	BM474441	BM474441	1022 bp	mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOURT_6490142 NIH_MGC_71	Homo sapiens	cdna clone	IMAGE:5521703
ACCESSION	5', mRNA sequence.	BM474441		
VERSION	BM474441			
KEYWORDS	BM474441.1	GI:18523483		
SOURCE	EST.			
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-re@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12187 row: 1 column: 24 High quality sequence stop: 675. Location/Qualifiers			
FEATURES	source			
	1..1022			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:5521703"			
	/clone_lib="NIH_MGC_71"			
	/tissue_type="leiomyosarcoma"			
	/lab_host="DH10B (phage-resistant)"			
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. "			
BASE COUNT	370 a 265 c 259 g 121 t	7 others		
ORIGIN				
Query Match	7.1%;	Score 691;	DB 10;	Length 1022;
Best Local Similarity	99.7%;	Pred. No. 1.3e-313;		
Matches 791;	Conservative	0;	Mismatches 2;	Indels 0;
Qy	7883	aaagcacacacacacaaagaaccagcagcggtgatgaggccatcaaatattgaac	7942	
Db	18	AAACACACACACACAAAGAACCCAGCGGTGATGAGGGCATCAAGATTATTGAAC	77	
Qy	7943	aacgtgcaagaagaacacaaacccacgtagaagaaggaaacccagcaggagaagccaag	8002	
Db	78	AACGTGCAAGAAAGAACCAACCAACCCAGTAGAGAGGAACCCAGCAGGAGAGGCCAAG	137	
Qy	8003	cacctaaagaaaagggcccaacccctggaagacctggccggttcacagagctctctgaa	8062	
Db	138	CACCTAAGAAAAGGCCAACCCCTGGAAGACTGGCGGCTTCACAGAGCTCTCTGAAA	197	

Qy	7068	aaagcaacgccc 7080 																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
----	------	-----------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Qy 8777 aaacaccaggccacactgaggaaactggcaaatgggtgctgctgatagctttcaaacgctc 8836
|||||
Db 361 AAACACCAGGCCACACTGAGGAACCTGGCAAAATGGTGTCTGTATAGCTTTTACAAGCGCTC 420

Qy 8837 caaagcaaacacctgacagtggaaaaacctctaaaaataatcagaagaagtcttcggggccc 8896
|||||
Db 421 CAAGCAAACACCTGACAGTGGAAACCTCTAAAAAATATCAGAAAGAGTCTTCGGGGCCC 480

Qy 8897 ctaagtagaacccgtggagacgtggtaagcaccagagaccctgtaaaaatcacaaaagca 8956
|||||
Db 481 CTAAGTAGAACCCGTGGGAGACGTGGTAAGCACAGAGACCCCTGTAAATACACAAAGCA 540

Qy 8957 aaagcaacacttcctgccccactgcccctcaagagggggggtggcgaagatggaagcg 9016
|||||
Db 541 AAAGCAACACTTCCCTGCCCTCCACTGCCCTTCAAGAGGGGAGGTGGCAAGATGGAAGCG 600

Qy 9017 tcaggggaaccaagaggtgcgtcgtcatgccagcacagagagaaattgtgaggagctgc 9076
|||||
Db 601 TCAGGGGAACCAAGAGGTGGCGCTGTCATGCCAGCACAGAGAGAAATGTGGAGGAGCTGC 660

Qy 9077 cagccagcaagaagcagaggggtgtctccaggggcaagaggcaaatcatccgaaccctgg 9136
|||||
Db 661 CAGCCAGCAAGACAGAGGGTGTCTCCAGGGCAAGAGGCAATCATCCGAACCCGTGG 720

Qy 9137 tcataatgaagaaagttgaggactctgcacaaagaattgaacctgcgga 9188
|||||
Db 721 TCATCATGAAGAGAAGTTTGAGGACTTCTGCAAAAAAAGTAATTGAACCTGCGGA 772

RESULT 6

BM455791
LOCUS BM455791 1070 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6412973 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498229
5', mRNA sequence.
ACCESSION BM455791
VERSION BM455791.1 GI:18504831
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1070)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaops-r@mail.nih.gov
Tissue Procurement: Lou Staudt

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12129 row: j column: 22
High quality sequence stop: 585.
Location/Qualifiers
1..1070
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5498229"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES source

1..1070
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5498229"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 386 a 291 c 238 g 155 t
ORIGIN

Query Match 5.9%; Score 576; DB 10; Length 1070;
Best Local Similarity 99.6%; Pred. No. 1.7e-259;
Matches 726; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5063 caagtctaaactggcagcagaagcggcgtgagaactcctaagggaaagtctgaagtcctcgtg 5122
|||||
Db 1 CAAGTCTTAACCTGGCAGCAAGAGGAGCTGAGAACTCTTAAAGGAAAGTCTGAAGTCCCTG 60

Qy 5123 aagacctggcggcgtctcatcgagctctccagacaccacaaagtcacactaaggaatcaatga 5182
|||||
Db 61 AAGACCTGGCCGGCTTCATCGAGCTTCCACACACCAAGTCACACTAAGGAATCAATGA 120

Qy 5183 ctaatgaaaaaactaccaaaagtatactacagagcttcacagccagacctagtggacacc 5242
|||||
Db 121 CTAACGAAAAAAATACCAAAGTATCCTACAGAGCTTCACAGCCAGACCTAGTGGACACC 180

Qy 5243 caacaagctccaagccacagcccaagagaagtcctcagggaaaagcagacactgaagaagaat 5302
|||||
Db 181 CAACAAGCTCCAAGCCACAGCCCAAGAGAAAGTCTCAGGAAAGCAGACACTGAAGAAAGAAT 240

Qy 5303 ttttagcttaggaaacaaacgcccacatcagcagcagcaagccatgcacacacccaaaccag 5362
|||||
Db 241 TTTTAGCATTTAGGAAACAAACGCCATCAGCAGGCAAAAGCCATGCACACCCCAACACCAG 300

Qy 5363 cagtaggtgaagaaagacatcaacacgctttttgggaactccagtcgagaaacttgacc 5422
|||||
Db 301 CAGTAGGTGAAGAGAAAGACATCAACACGTTTTTGGGAACCTCCAGTGCAGAACTGGACC 360

Qy 5423 agccaggaaatttaacctggcagcaatagacggctacaaactcgtgaaggaaagccaggg 5482
|||||
Db 361 AGCCAGGAAATTTAGCTGGCAGCAATAGACGGCTACAAAATCTGTAAGGAAAAAGGCCAGG 420

Qy 5483 cctcagaagaactgactggcttcagagactttccagacacatgcactgataacccca 5542
|||||
Db 421 CTCTAGAAGAACTGACTGGCTTCAGAGAGCTTTTCCAGACACATGCATGATAACCCCA 480

Qy 5543 cagctgatgagaaaaactaccaaaaaataactctgcaaatctccgcaatcagaccccgagg 5602
|||||
Db 481 CGACTGATGAGAAAAACTACCAAAAAAATACTCTGCAAAATCTCCGCAATCAGACCCAGCG 540

Qy 5603 acaccccaacaaacacaaagcaacgcccgaagagagcctcaagaagcagacgtagagg 5662
|||||
Db 541 ACACCCCAACAAACAAACAAAGCAACGCCCAAGAGAGCCTCAAGAAAGCAGACGTAGAGG 600

Qy 5663 aagaatttttagcattcaggaactaaacacacatcagcagcagcaagccatgcacacgctca 5722
|||||
Db 601 AAGAATTTTACATTCAGGAAACTAACACCATCAGCAGGCAAAAGCCATGCACACCCCTA 660

Qy 5723 aagcagcagtagtgaaagagaaagacatcaacacatttgtgggactccagtgaggagaaac 5782
|||||
Db 661 AAGCAGCAGTAGTGAAGAGAAAGACATCAACACATTTGTGGGGACTCCAGTGGAGAAAC 720

Qy 5783 tggacctgc 5791
|||||
Db 721 TGGACCTGC 729

RESULT 7
BG403181
LOCUS BG403181 697 bp mRNA linear EST 12-MAR-2001
DEFINITION 602418615F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4525994 5',
mRNA sequence.
ACCESSION BG403181
VERSION BG403181.1 GI:13296629
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 697)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10432 row: m column: 03
High quality sequence stop: 688.
Location/Qualifiers

FEATURES
source

1. .697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4525994"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
263 a 182 c 163 g 89 t

BASE COUNT
ORIGIN

Query Match 5.9%; Score 575; DB 10; Length 697;
Best Local Similarity 100.0%; Pred. No. 4.6e-259;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7734 tgacaaacacacaaaattccctgcaaatctcccccacagaaactaacagacactgccac 7793
Db 108 TGACAAAACACAAAAATTCCTGCAAAATCTCCCCACACAGAACTAACAGACACTGCCAC 167
QY 7794 gacacaaagagatgccccaaagacacgtcccgaggaaagtaaaagagagctctcagc 7853
Db 168 GAGCACAAAAGAGATGCCCAAGACAGCTCCAGGAAGAAGTAAAGAGAGAGCTCTCAGC 227
QY 7854 agttgagagctcacgcaaacatcaggcgaagcaaacgacacacacacacacacacgaag 7913
Db 228 AGTTGAGAGGCTCAGCGAAACATCAGGGGCAAGACACACACACAAAAGAACCAAG 287
QY 7914 cgtgatgagggcatcaaatattgaagcaactgtgcaagaaagaaacccacgtaga 7973
Db 288 CGGTGATAGGGCATCAAGATTGAAGCAACGTGCAAGAGAAGAAACCAACCCAGTAGA 347
QY 7974 agaggaaacccagcagagagaagccagagacactaaggaaaaggcccaacccctggaaag 8033
Db 348 AGAGGAACCCAGCAGGAGAAGGCGCAAGAGCACCCTAAGGAAAAGGCCCAACCCCTGGAAGA 407
QY 8034 cctggccgcttcacagagactctgaaacatcagtcacactcaggaatcactgactgc 8093
Db 408 CTTGGCCGGCTTCACAGAGCTCTCTGAACATCAGGTACACTCAGGAATCAGTACTGCTC 467
QY 8094 tggcaagccactaaaataccctgcgaattctccccactagaaagtgtgtagacaccacagc 8153
Db 468 TGCAAAAGCCACTAAATACCTCGCAATCTCCCCACTAGAAAGTGGTAGACACACACAGC 527
QY 8154 aagcacaagagagcatctcaggacacgctgtgcagaagggtacaaagtaaaagagagccttc 8213
Db 528 AAGCACAAAAGAGGCATCTCAGGACAGCTGTGCAGAGAAGGTACAAGTAAAAAGAGAGCCTTC 587
QY 8214 aqcagtcaggttcacacaacatcagggaacccacgagtagacacacacacacacacag 8273
Db 588 AGGAGTCAAGTTTACACAAACATCAGGGGAAACACCGGATGACAGACAAAAGACCAAGCAGG 647
QY 8274 tgaagtaagggcatcaaaagcattgaaggaattctg 8308
Db 648 TGAAGATAAAGGCATCAAGGCAATTGAAGGAATCTG 682

RESULT 8

BM452654
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT_6401184 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498641
BM452654
ACCESSION
VERSION
BM452654.1 GI:18501694
KEYWORDS
EST.
SOURCE
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1068)
NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-re@mail.nih.gov

TITLE

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

JOURNAL

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12130 row: l column: 02
High quality sequence stop: 535.
Location/Qualifiers

FEATURES

source

1. .1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5498641"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
376 a 290 c 253 g 147 t 2 others

BASE COUNT
ORIGIN

Query Match 5.7%; Score 553; DB 10; Length 1068;
Best Local Similarity 99.7%; Pred. No. 1.1e-248;
Matches 653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5546 ctgatgagaaaactcaccacaaaataactctgcaaatctccgcaatcagaccccgagaca 5605
Db 45 CTGATCAGAAAACCTACCAAAAAAATACTCTGCAAAATCTCGCAATCAGACCCAGCGACA 104
QY 5606 ccccaacaaacaaagcaacggcccaagaagaagctcagaagaagcagacgtagagaag 5665
Db 105 CCCCAACAAACACAAAGACACAGCCCCAAGAGAGCCCTCAAGAAAGCAGACGTAGAGGAAG 164
QY 5666 aattttagcatcaggaactaacaccatcagcaggcaaaagcctagcacgcctctaaag 5725
Db 165 AATTTTACCATTCAGGAAACTAACCATCATCAGCGCAAGCCATGCACACCCCTTAAG 224
QY 5726 cagcagtaggtgaagagaagacatcaacacatttgtggggactccagtggaagaaactgg 5785
Db 225 CAGCAGTAGGTGAAGAGAAGACATCAACACATTTGTGGGGACTCCAGTGGAGAACTGG 284
QY 5786 acctgcagaagaaatttacctggcagcagaagacggccacacaaactcctaaagaaagccca 5845
Db 285 ACTGCTAGGAAATTTACCTGGCAGCAAGAGACGGCCACAACTCTCTAAAGAAAGGCCCA 344
QY 5846 aggccttagaagatctggctggcttcaagagagctctccagacacacaggttcacactgagg 5905
Db 345 AGGCTCTAGAAGATCTGGCTGGCTTCAAGAGAGCTCTTCCAGACACCCAGGTCACACTGAGG 404


```
Qy 5906 aatcaatgaccgatgacaaaatcacagaagtatctctgcaaatctccacaaccagaccag 5965
Dy 405 AATCAATGACCGATGACAAAATCACAGAAGTATCTCTGCAAAATCTCCACAACGACCCAG 464
Qy 5966 tcaaaaccccaaacgctccaagcaacgactcaagatatcttggaagtagtgatga 6025
Dy 465 TCAAAACCCCAACAGCTCCANGCAACGACTCAAGATATCTTGGGAAAGTAGGTGGA 524
Qy 6026 aagaaggggtctaccagtggcgaagctcacacagactgaggaagaccacacagacac 6085
Dy 525 AAGAAGAGGTCTTACCAGTGGCAAGCTCACACAGAGCTCAGGGAAGACACACAGACAC 584
Qy 6086 acagagagacagcggggagtggaagagagcatcaaaagcgtttaaggaatctgcgaagcaga 6145
Dy 585 ACAGAGAGACAGCAGGAGATGGGAAGAGCATCAAAAGGCTTTAAGGAATCTGCAAAAGCAGA 644
Qy 6146 tgcctggaccagcaaacataggaactggatggagagtgagtgcccaagacacctaa 6200
Dy 645 TGCTGGACCCAGCAACATATGGAATCTGGATGGAGAGGTGGCCCAAGAACACCTAA 699

RESULT 9
LOCUS BE395002 606 bp mRNA linear EST 21-JUL-2000
DEFINITION 601308294F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3626638 5',
mRNA sequence.
ACCESSION BE395002
VERSION BE395002.1 GI:9340367
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM308 row: k column: 23
High quality sequence stop: 602.
FEATURES
source location/Qualifiers
1..606
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3626638"
/clone_lib="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 231 a 158 c 144 g 73 t

ORIGIN

Query Match 5.7%; Score 552; DB 10; Length 606;

Best Local Similarity 99.8%; Pred. No. 3e-248;

Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 7901 aagaacacagcagcggtgagggcatcaaaagtattgaagcaacgtagcaagaagaac 7960
|||||
```

```
Dy 1 AAGAACCCAGCAAGCGGTGATGAGGGCATCAAAAGTATTGAAGCAACGTCACAAAGAGAACAAC 60
Qy 7961 caaacccagtagaagagaaacccagcagagagaagcccaagcagcactaaggaagagcc 8020
Dy 61 CAACCCAGTAGAGAGAGAACCCAGCAGAGAGAGCCCAAGAGCAGCTTAAGGAAAAGGCC 120
Qy 8021 aacccctggaagacctggccggtctcacagagctctctgaacatcaggtcacactcagg 8080
Dy 121 AACCCCTGGAAGAGACTGCGCGGTTCACAGAGCTCTCTGAAAACATCAGGTACACTCAGG 180
Qy 8081 aatcacactgactgctggcaaaagccactaaaataccctcggaatctccccactagaaagg 8140
Dy 181 AATCACTGACTGCTGGCAAAAGCCACTAAAATACCCCTGGAATCTCCCACTAGAAAGTGG 240
Qy 8141 tagacaccacagcagcaagcacaagagcgtctcagagacacgctgtgcagaaggacaaagtaa 8200
Dy 241 TAGACACCACAGCAAGCACAAGAGGCATCTCAGGACACGCTGTGCAGAAGGTACAAAGTAA 300
Qy 8201 aagaagagccttcagcagcagtcgaagtccacacaacatcaggggaaacccagcagatgcagaca 8260
Dy 301 AAGAAGAGCCTTCAGCAGCTCAAGTTTCACACAACATCAGGGGAAACACCGATGCAGACA 360
Qy 8261 aagaacacagcaggtgaaagataaaaggcatcaaaagcattgaaggaatctgcaaaacagacac 8320
Dy 361 AAGAACCAGCAGAGGTGAAGATAAAGGCATCAAAAGCATTTGAAGGAATCTGCAAAACAGACAC 420
Qy 8321 cggctccagcagcagtgtaactggcagcagagagacgcccagagcaccacaggggaaagtg 8380
Dy 421 CGGCTCCAGCAGCAAGTGTAACTGGCAGCAGGACACGCCAAGAGACCCAGGGGAAGTGG 480
Qy 8381 cccaagccatagaagacacacacacacacacacacacacacacacacacacacacacacacac 8440
Dy 481 CCCAAGCCATAGAAGACACCTAGCTGGCTTCAAAAGACCCAGCAGCAGGTACACATGAAGAAT 540
Qy 8441 caatgactgatgacaaaaacacacacacacacacacacacacacacacacacacacacacacac 8500
Dy 541 CAATGACTGATGACAAAAACCACTAAATACCCCTGCAAAATCATCCACAGAACTAGAAGACA 600
Qy 8501 ccg 8503
|||||
Dy 601 CCG 603

RESULT 10
LOCUS AW835684/c 580 bp mRNA linear EST 18-MAY-2000
DEFINITION QVA-LT0016-140200-105-h06 LT0016 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW835684
VERSION AW835684.1 GI:7929658
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 580)
Dias Netto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
```

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV4-LT0016-140
200-105-h06t3-2000-02-14&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 577.
Location/Qualifiers

FEATURES

source
1..580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LT0016"
/dev_stage="Adult"
/note="Organ: leiomyos; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 83 a 140 c 147 g 209 t 1 others
ORIGIN

Query Match 5.6%; Score 544; DB 9; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.7e-244;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5687 taacacatcagcagcaagccatcacacgcctcaagcagcagtagtggaagaaaag 5746
|||||
Db 579 TAACACCATCAGCAGGCAAGCCATGCACAGCGCTTAAGCAGCAGTAGGTGAAGAAAAG 520
QY 5747 acatcaacacatttggggactccagtggagaaactggacctgctaggaattttacctg 5806
|||||
Db 519 ACATCAACACATTTGGGGACTCCAGTGGAGAAACTGGACCTGCTAGGAATTTACCTG 460
QY 5807 gcagcaagagcggccacaaaactcctaaagaaaagcccaagcgtctagaagatcttgctg 5866
|||||
Db 459 GCAGCAAGAGACGGCCACAAAACCTCTTAAAGAAAGGCCAAGGCTCTAGAAGATCTGGCTG 400
QY 5867 gctcaagagctctccagacacacaggtcacactgaggaatcaatgaccgatgacaaaa 5926
|||||
Db 399 GCTTCAAGAGCTCTTCCAGACACCAAGGTCACACTGAGGAATCAATGACCGATGACAAA 340
QY 5927 tcacagaagtatactgcaaatctccacacagaccagccagtcacaaaccccaacagctcca 5986
|||||
Db 339 TCACAGAAGTATCTGCAAAATCTCCACACACGACCCAGCTCAAAACCCCAACAAAGCTCCA 280
QY 5987 agcaagactcaagatatccttgggaaagttaggtgtgaaagaagaggttctaccagtctg 6046
|||||
Db 279 AGCAACGACTCAAGATATCTTGGGAAAGTAGGTGTGAAGAAGAGGTCTTACCAGTGG 220
QY 6047 gcaagctcacagcagctcagggaaagaccacacagacacacagagagacagcagagatg 6106
|||||
Db 219 GCNAGCTCACACAGACTCAGGGAAGACCCACACAGACACACAGAGAGACGACGAGGAGATG 160
QY 6107 gaaagagcatcaagcgtttagggaatctgcaaaagcagatgctggacccacgacaaactatg 6166
|||||
Db 159 GAAAGAGCATCAAAAGCTTTAAGGAATCTGCAAAAGCAGATGCTGGACCCAGCAAAACTATG 100
QY 6167 gaactgggtgagagaggtggccaagaacaccttaagaagaagggcccaatcactagaagacc 6226
|||||
Db 99 GAACTGGGATGGAGAGGTGGCCAAAGAACACCTAAGGAAGAGGGCCCAATCACTAGAGAGCC 40
QY 6227 tgcc 6230
|||||
Db 39 TGCG 36
RESULT 11
BM476598
LOCUS
DEFINITION AGENCOURT_6476446 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5553546
1047 bp mRNA linear EST 05-PEB-2002

5', mRNA sequence.
BM476598
BM476598.1 GI:18525640
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1047)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12370 row: k column: 19
High quality sequence stop: 608.
Location/Qualifiers
1..1047
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5553546"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SpORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

BASE COUNT 372 a 281 c 244 g 149 t 1 others
ORIGIN
Query Match 5.5%; Score 537; DB 10; Length 1047;
Best Local Similarity 99.5%; Pred. No. 3.8e-241;
Matches 737; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5245 acaagctcaagccagcccaagagaagctctcaggaaagcagacactgaagagattt 5304
Db 1 ACAAGTCCACAGCCAGCCCAAGAGAAGTCTCAGGAAGACAGACTGAAGAAGAAATTT 60
QY 5305 tttagcatttagaatacaaacgccaatcagcaggcaaaagccatgcacacaccccaaccagca 5364
Db 61 TTAGCATTTTAGGAAACAAACGCCATCAGCAGGCAAGCCATGCACACACCCAAACCCAGCA 120
QY 5365 gtaggtagaagaaagacatcacacgtttttgggaactccagtcagaaactggaccag 5424
Db 121 GTAGGTGAAGAGAAACACATCAACACGTTTTTTGGGAACCTCCAGTGCAGAACTGGACCAG 180
QY 5425 ccagggaatttacctggcagcaatagacgctacaactcgttaaggaaaagcccaggct 5484
Db 181 CCAGGAATTTACTTGGCAGCAATAGCAGGCTACAACTCCTTAAGGAAAGGCCAGGCT 240
QY 5485 ctagaagaactgactggcttcagagagctttccagacaccatgcactgataaccacca 5544
Db 241 CTAGAAGAAGTCTGCTTTCAGAGAGCTTTTCCAGACACCATGCATGATATACCCACG 300
QY 5545 gctgatgagaaaactacaaaaataactctgcaaatctccgaatccagacccagcggac 5604
Db 301 ACTGATGAGAAAACCTACCAAAAAAATACTCTGCAAAATCTCCGCAATCAGACCCAGCGGAC 360
QY 5605 acccaacaaacacaaagcaacgcccagagagaagcctcaagaagcagacgtagaggaa 5664
Db 361 ACCCCAAACAAACAAAGCAACAGCCCAAGAGAGGCTCAAGAAAGACAGACGCTAGAGGAA 420
QY 5665 gaatttttagcttcagggaactaacacacccatcagcaggcaagccatgcacacgcctaaa 5724

```
|||||
Db 421 GAAATTTTAGCATTCAGAAATTAACACCATCAGCAGCAAGCCATCCACACGCTATAA 480
Qy 5725 gcagcagtagtggaagaaagacatacaacacatttggggactccagtgagaaactg 5784
Db 481 GCAGCAGTAGTGGAAGAAAGACATCAACACATTTGTGGGGCTCCAGTGGAGAACTG 540
Qy 5785 gacctgctaggaatttacctggcagcaagagagcgccacaaaactcctaaagaaagccc 5844
Db 541 GACCTGCTAGGAAATTTACCTGGCAGCAAGAGACGGCCACAAACTCCTAAGAAAGGCC 600
Qy 5845 aagcctctagaaagtctggcttcaaaagagctctccagacaacacaggtcaactgag 5904
Db 601 AAGCTCTAGAAAGATCTGGCTGGCTTCAAGAGCTCTTCCAGACACACAGGTCCACTGAG 660
Qy 5905 gaatacaatgaccgatgacaaatcacagaagtatcctgcaaatctccacaacccagccca 5964
Db 661 GAATCAATGACCGATGACCAATACAGAGATATCTTGCAAAATCTCCACAACGACCCA 720
Qy 5965 gtcaaaaccccaacaagctcc 5985
Db 721 GTCAAAACCCCAACAAAGCTCC 741

RESULT 12
AUI40843
LOCUS AUI40843 PLACE4 Homo sapiens cDNA clone PLACE4000347 5', mRNA
DEFINITION AUI40843 822 bp mRNA linear EST 25-OCT-2000
sequence.
ACCESSION AUI40843
VERSION AUI40843.1 GI:11002364
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,Y., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isoqai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,Y., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isoqai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isoqai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
location/Qualifiers
1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE4000347"
/clone_lib="PLACE4"
/tissue_type="placenta"
/note="vector: pME18SFL3"
BASE COUNT 305 a 211 c 177 g 125 t 4 others
ORIGIN
Query Match 5.5%; Score 535; DB 9; Length 822;
Best Local Similarity 99.6%; Pred. No. 3.2e-240;
Matches 685; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4941 agcagttggaagctcacagacatcaggagagactacacacacacacagagcccaac 5000
Db 1 AGCAGTTGGCAAGCTCACAGACATCAGGAGAGACTACACACACACAGACGCCAAC 60
```

```
Qy 5001 aggagatgtagagcatgaaagcatttatggagtgctccaaagcagatcttagactcagc 5060
Db 61 AGGAGATGGTAAGACCATGAAGCATTTATGGAGTCTCCAAAGCAGATCTTAGACTCAGC 120
Qy 5061 agcaagtctaactggcagcaagagcagctgagaactcctaagggaagtctgaagtccc 5120
Db 121 AGCAAGTCTAATGTGCAGCAAGAGCGCAGTGAGAACTCTCTAAGGGAAAGTCTGAAGTCCC 180
Qy 5121 tgaagacctggccggcttcacagctcttcacagacacaaagtcacactaagaatacaat 5180
Db 181 TGAAGACCTGGCGCGCTTCATCGAGCTCTTCAGACACACCAAGTCACACTAAGAGATTCAAT 240
Qy 5181 gactaatgaaaaaactcacaagatcctacagagcttcacagcagacacctagtgtgacac 5240
Db 241 GACTTAACGAATAAACTACCAAGATATCTTACAGAGCTTCACAGCCAGACCTAGTGGACAC 300
Qy 5241 cccaacagctccagccagccacagcccaagagagtgctcagggaagcagacactgaagaaga 5300
Db 301 CCCAACAAAGCTCCAAGCCACAGCCCAAGAGAAAGTCTCAGGAAAGCAGACACTGAAGAAGA 360
Qy 5301 atttttagcatttagaaaacaaacccatcagcagcagcaagccatgcacacaccccaaac 5360
Db 361 ATTTTTCAGATTTTGAAGAACAAACGCCATCAGCAGGCAAGCCATGCACACACCCCAACCC 420
Qy 5361 agcagtagtgtaagagaagacatcaacacgctttttgggaactccagtgacagaaactgga 5420
Db 421 AGCAGTAGGTGAAGAGAAAGACATCAACACGTTTGTGGAACTCCAGTGCAGAAACTGGA 480
Qy 5421 ccagccaggaatttacctgagcagcaatagacggtcacaactctgaaggaaggaagccca 5480
Db 481 CCAGCCAGGAAATTTACCTGGCAGCAATAGACGGGTACAAACTCGTAAAGGAAAGGCCCA 540
Qy 5481 ggctctaagaaactgactggcttcagagagctttccagacacaccatgcactgataaacc 5540
Db 541 GGCTCTAGAAGAACTTGACTGGCTTCAGAGAGCTTTTCCAGACACCATGCACGTGATAAACC 600
Qy 5541 cacagctgatgagaaactaccaaaatactctgcaaatctccgaatccgcaatcagaccagc 5600
Db 601 CACAACTGATGAGAAACTACCAAAAAATACTCTGCAAAATCTCCGCAATCAGACCCAGC 660
Qy 5601 ggacaccccccaacacacaaagcaacgg 5628
Db 661 GGACACCCCAACAAACAAAGCAACGG 688

RESULT 13
BI868409
LOCUS 603392387F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402684 5',
DEFINITION mRNA sequence.
ACCESSION BI868409
VERSION BI868409.1 GI:16042082
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 753)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12026 row: m column: 21
High quality sequence stop: 741.
```

FEATURES		Location/Qualifiers	
source		1..753	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5402684"	
		/clone_lib="NIH_MGC_90"	
		/tissue_type="adenocarcinoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."	
BASE COUNT		286 a 144 c 192 g 131 t	
ORIGIN			
Query Match		5.4%; Score 525; DB 10; Length 753;	
Best Local Similarity		99.8%; Pred. No. 1.6e-235;	
Matches 575; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
QY	9161	ctctcgaaagaattgaacctgcggaagagctgaacagcaacgacatgaaaacacaa	9220
Db	13	CTTCTCGAAAAGAAATTGAACCTGCGAGAGAGCTGAACAGCAACGACATGA	72
QY	9221	aagaggaaacacaaattacaaagactcggtccctgaaataaagggaatatccctgcgctcca	9280
Db	73	AAGAGGAACACAATTTACAAGACTCGTGCTCCCTGAAATAAAGGGAATATCCTGCGCTCCA	132
QY	9281	gacgccaaagataagactgaggagagaacagacaataactgaggtctttgtattagcagaaa	9340
Db	133	GAGCGCAAAATAAGACTGAGGCGAAGAACAGACAAATACTGAGGTCTTTGTATTAGCAGAAA	192
QY	9341	gaatagaatacaaacagaaatgaaagaagcccatgaagacctccccagagatggacattc	9400
Db	193	GAATAGAAATAACAGAAATGAANAAGCCCATGAAGACCTCCCCAGAGATGGACATTCC	252
QY	9401	agaatccagatgatggagcccgaaacccataccttagagacaaaagtcaactgagaacaaaa	9460
Db	253	AGAATCCAGATGATGAGCCCGGAAACCCATACCTAGAGACAAAGTCACTGAGAACAAAA	312
QY	9461	ggctgtgaggctgtgtagacagaataagagctcccagacctaaagtgaggagagagcg	9520
Db	313	GGTGTGTTGAGGTCTGTAGACAGAAATGAGAGCTCCAGAGCTTAAGGTGGCAGAGGAGCG	372
QY	9521	gagggcagaagagtcggaaggttctcatgcagaatcagaaagggaagagagcaggaa	9580
Db	373	GAGGGCAGAAAGATGCGAAGTTCTCATGCAGANTCAGAAAGGGAAGGAGAGACAGGAA	432
QY	9581	attcagactccatgtgcctgagatcaagaaagacaaaaagcccgctgcagcaagcactt	9640
Db	433	ATTCAAGACTCCATGTGCCTGAGATCAAGAAAGACAAAAAGCCAGCCTGCAGCAAGCACTT	492
QY	9641	tggagagcaaatctgtcagagagataacgcggagtgctcaagaggtgtgcagaaaaatccaa	9700
Db	493	TGGAGAGCAAAATCTGTGCAGAGAGTAACCGGAGTGTCAAGAGGTGTGCAGAAAAATCCAA	552
QY	9701	agagagctgaggacaatgtgtgttcaagaaaaataa	9736
Db	553	AGNAGGCTGAGGACATGTGTGTCAAGAAATAA	588
RESULT 14		751 bp mRNA linear EST 13-FEB-2001	
BG252689		LOCUS	
DEFINITION		602366335F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4474376 5', mRNA sequence.	
ACCESSION		BG252689	
VERSION		BG252689.1	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		1 (bases 1 to 751)	
TITLE		NIH-MGC http://mgc.nci.nih.gov/.	
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT		Unpublished (1999)	
		Contact: Robert Strausberg, Ph.D.	
		Email: cgapbs-re@mail.nih.gov	
		Tissue Procurement: ATCC	
		cDNA Library Preparation: Life Technologies, Inc.	
		DNA Sequencing by: Incyte Genomics, Inc.	
		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
		Plate: LLAM10298 row: f column: 09	
		High quality sequence stop: 667.	
FEATURES		Location/Qualifiers	
source		1..751	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4474376"	
		/clone_lib="NIH_MGC_90"	
		/tissue_type="adenocarcinoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."	
BASE COUNT		273 a 198 c 171 g 109 t	
ORIGIN			
Query Match		5.3%; Score 519; DB 10; Length 751;	
Best Local Similarity		99.7%; Pred. No. 1.1e-232;	
Matches 619; Conservative		0; Mismatches 2; Indels 0; Gaps 0;	
QY	6642	gaaactacacaaatagcctgcagatctccacacacagaccctagtggtaccaccaaat	6701
Db	1	GAAACTACCAAAATAGCCTGCAGATCTCCACAACAGACCCAGTGGGTACCCCAACAAT	60
QY	6702	cttcaggccacagtcacagagaagtcctcaggaaagcagacgtagagaagaatccttagc	6761
Db	61	CTTCAAGCCACAGTCCAGAGAGTCTCAGGAAGCAGATGTAGAGAGAAATCCTTAGC	120
QY	6762	actcaggaacgaacaccatcagtaggaaagctatggacacacccacacacagcaggag	6821
Db	121	ACTCAGAAACGACACCATCAGTAGGGAAAGCTATGGACACACCCCAACCCAGCAGG	180
QY	6822	tgatgagaagacatgaaagcatttatgggaactccagtcagaaattggacctgccagg	6881
Db	181	TGATGAGAAAGACATCAAGCATTATTTGGAACTCCAGTGCAGAAATTTGACCTGCCAG	240
QY	6882	aaatttacctggcagcaaaagatggccacaaactcctaaggaaaaagccaggctctaga	6941
Db	241	AAATTTACCTGGCAGCAAAAGATGGCCCAAACTCCTTAAGGAAAAGGCCAGGCTCTAG	300
QY	6942	agacctggctggcttcaagagcctctccagacacacagcagcactgacaaagccacgactga	7001
Db	301	AGACCTGGCTGGCTTCAAGAGCTCTTCCAGACACCCAGGCACCTGACACGCCACGACTGA	360
QY	7002	tgagaaactacaaaatagcctgcaaatctccacacacagaccctagtgagacccccagc	7061
Db	361	TGAGAAAACTACCAAAATAGCCTGCAAAATCTCCACAACAGACCCAGTGGACACCCAGC	420
QY	7062	aggcacaaagcaacggcccaagaacacacctcaggaaagcagacgtagagagaatctttt	7121
Db	421	AAGSCAAAAGCAACGCCCCCAAGAGAAACCTCAGGAAAGCAGACGTAGAGAGAAATTTT	480
QY	7122	agcactcaggaaacgaacaccatcagcagggcaagcctatggacacaccccaaaaacagcag	7181
Db	481	AGCACTCAGAAACGACACCATCAGCAGGCAAGCCATGGACACACACCAAAACCCAGCT	540

QY 7182 aagtatgagaaaaatatacaacacatttgtggaactccagtgcagaaactggacctgct 7241
|||||
Db 541 AAGTGATGAGAAAAATATCAACACATTTTGTGGAACCTCCAGTGCAGAAACTGGACCTGCT 600
QY 7242 aggaatttacctgcagcaa 7262
|||||
Db 601 AGGAAATTACCTGGCAGCAA 621

RESULT 15
BG677474
LOCUS BG677474 692 bp mRNA linear EST 01-MAY-2001
DEFINITION 602625107f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4750249 5',
mRNA sequence.
ACCESSION BG677474
VERSION BG677474.1 GI:13908871
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM10604 row: m column: 02
High quality sequence stop: 690.
Location/Qualifiers
1..692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4750249"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 247 a 170 c 174 g 101 t
ORIGIN

Query Match 5.2%; Score 510; DB 10; Length 692;
Best Local Similarity 100.0%; Pred. No. 1.8e-228; Mismatches 0; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8752 ctggccagcttccaagagctctctcaaacaccagccacactgaggaactggcaaatggt 8811
|||||
Db 40 CTGGCCAGCTTCCAAGAGCTCTCTCAACACCAGCGCCACACTGAGGAACGTGCAATGGT 99

QY 8812 gctgctgtagctttacaagcgctccaagaacacacacctgacagtggaaaacctctaaaa 8871
|||||
Db 100 GCTGCTGATAGCTTTACAAAGCGCTCTCAAAAGCAAAACACCTGACAGTGGAAAACTCTAAAA 159

QY 8872 atatccagaagaagttctcgggcccctaaagttagaacccgtgggagacgtggttaagcacc 8931
|||||
Db 160 ATATCCAGAAGAGTTCCTTCGGGCCCTTAAAGTAGAACCCCGTGGGAGACGTGGTAAGCACC 219

QY 8932 agagacctgtaaaaatcaaaagcaaacacttccctgcgcccccactgcctctcaag 8991
|||||
Db 220 AGAGACCTGTAAATATCAAAAGCAAAAGCAACACTTCCCTGCCCCCACTGCCCTTCAAG 279

QY 8992 agggaggtggaagatggaagcgctacggaaccaagaggtgcgtgcatgccagca 9051
|||||

Db 280 AGGGGAGGTGGCAAAAGATGGAAGCGCTACGGGAACCAAGAGGCTGCGCTGCATGCCAGCA 339
QY 9052 ccagaggaatttgtagagagctgccagccagcaagaagacagaggttgcctccagggca 9111
|||||
Db 340 CCAGAGGAAATTTGTGGAGGAGCTGCCAGGCCAGCAAGAGAGGGTTGCTCCAGGGCA 399
QY 9112 agaggcaaatcatccgaaccccggtggtcatcatgaagaagaagtttgaggacttctgcaaaa 9171
|||||
Db 400 AGAGGCAAAATCATCCGAACCCGCTGGTTCATCATGAAGAGAAGTTTGAGGACTTCTGCAAAA 459
QY 9172 agaattgaacctgcggaagagctgaacgaacgacatgaaacccaacaaagagggaacac 9231
|||||
Db 460 AGAATTGAACTTCCGGAAGAGCTGAACGAACGACATGAAAAACCAACAAAGAGGAACAC 519
QY 9232 aaattacaagactcggctccctgaaaaataag 9261
|||||
Db 520 AATTACAAGACTCGGTCCCTGAAATAAG 549

Search completed: August 23, 2002, 20:25:27
Job time: 9627 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 251.89 Seconds
(without alignments)
9523.430 Million cell updates/sec

Title: US-09-700-906A-1_COPY_197_9962
Perfect score: 9766
Sequence: 1 atgtggccacgacgcct.....tcatagggacagtgaagata 9766

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/lna/5A-COMB.seq.*
2: /cgn2_6/ptodata/2/lna/5B-COMB.seq.*
3: /cgn2_6/ptodata/2/lna/6A-COMB.seq.*
4: /cgn2_6/ptodata/2/lna/6B-COMB.seq.*
5: /cgn2_6/ptodata/2/lna/PCPUS-COMB.seq.*
6: /cgn2_6/ptodata/2/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	389	4.0		561	4	US-09-328-111-458	Sequence 458, Appl
C 2	20	0.2		340	1	US-08-599-232-105	Sequence 105, Appl
C 3	20	0.2		340	5	PCT-US96-06352-105	Sequence 105, Appl
C 4	20	0.2		340	5	PCT-US96-06583-105	Sequence 105, Appl
C 5	20	0.2		530	4	US-08-991-789A-221	Sequence 221, Appl
C 6	20	0.2		530	4	US-09-062-451-221	Sequence 221, Appl
7	20	0.2		7881	2	US-08-751-189-1	Sequence 1, Appl
8	20	0.2		7881	2	US-09-060-836-1	Sequence 1, Appl
9	20	0.2		7881	4	US-09-184-445-1	Sequence 1, Appl
10	19	0.2		259	1	US-08-222-177A-24	Sequence 24, Appl
11	19	0.2		1325	4	US-08-915-795-6	Sequence 6, Appl
12	19	0.2		2040	1	US-08-599-252-103	Sequence 103, Appl
13	19	0.2		2040	5	PCT-US96-06352-103	Sequence 103, Appl
14	19	0.2		2040	5	PCT-US96-06583-103	Sequence 103, Appl
C 15	19	0.2		2169	3	US-08-806-326-5	Sequence 5, Appl
C 16	19	0.2		3286	4	US-09-211-417-2	Sequence 2, Appl
C 17	19	0.2		8797	2	US-08-723-306-6	Sequence 6, Appl
C 18	19	0.2		8797	5	PCT-US96-10041-6	Sequence 6, Appl
C 19	19	0.2		11093	2	US-08-723-306-5	Sequence 6, Appl
C 20	19	0.2		11093	5	PCT-US96-10041-5	Sequence 5, Appl
C 21	19	0.2		17056	4	US-09-245-041-3	Sequence 3, Appl
C 22	18	0.2		20	4	US-09-593-589-35	Sequence 35, Appl
C 23	18	0.2		47	1	US-08-222-177A-122	Sequence 122, Appl
24	18	0.2		47	1	US-08-222-177A-409	Sequence 409, Appl
25	18	0.2		50	1	US-08-222-177A-328	Sequence 328, Appl
26	18	0.2		51	1	US-08-222-177A-424	Sequence 424, Appl
27	18	0.2		65	1	US-08-222-177A-116	Sequence 116, Appl

28	18	0.2	65	1	US-08-222-177A-421	Sequence 421, App
29	18	0.2	72	1	US-08-222-177A-131	Sequence 131, App
30	18	0.2	72	1	US-08-222-177A-427	Sequence 427, App
31	18	0.2	88	3	US-08-545-860D-93	Sequence 93, Appl
32	18	0.2	109	3	US-08-532-896-29	Sequence 29, Appl
33	18	0.2	186	1	US-08-222-177A-12	Sequence 12, Appl
34	18	0.2	194	1	US-08-222-177A-15	Sequence 15, Appl
35	18	0.2	225	3	US-08-545-860D-91	Sequence 91, Appl
36	18	0.2	228	3	US-08-545-860D-56	Sequence 56, Appl
37	18	0.2	228	5	PCT-US94-04496-56	Sequence 56, Appl
38	18	0.2	240	1	US-08-222-177A-30	Sequence 30, Appl
39	18	0.2	264	1	US-08-222-177A-10	Sequence 10, Appl
40	18	0.2	288	4	US-09-556-868-1	Sequence 1, Appl
41	18	0.2	298	1	US-08-599-252-88	Sequence 88, Appl
42	18	0.2	298	5	PCT-US96-06352-88	Sequence 88, Appl
43	18	0.2	298	5	PCT-US96-06583-88	Sequence 88, Appl
44	18	0.2	354	1	US-07-953-230A-7	Sequence 7, Appl
45	18	0.2	439	1	US-08-320-559-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-328-111-458
; Sequence 458, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertli, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 458
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(561)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-458

Query Match 4.0%; Score 389; DB 4; Length 561;
Best Local Similarity 100.0%; Pred. No. 3.5e-167;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	6691	acccacaacatttcaagccacagtcacagagaagtcctcaggaaagacagcgtagaggaa	6750
Db	1	acccacaacatttcaagccacagtcacagagaagtcctcaggaaagacagcgtagaggaa	60
QY	6751	gaatccttagcactcaggaacacacacatcagtaggagaaagctatgacacaccaa	6810
Db	61	gaatccttagcactcaggaacacacacatcagtaggagaaagctatgacacaccaa	120
QY	6811	ccagcaggaggatgagagaagacatgaaagcatttatgggaactccagtcagaaattg	6870


```
Db 121 cccgagggaggtgatgagaagcatgaaagcatttatgggaactccagtgcagaaattg 180
QY 6871 gacctgcaggaatttacctggcgagcaaaagatggccacaaactctctaaagaaaggcc 6930
Db 181 gacctgcaggaatttacctggcgagcaaaagatggccacaaactctctaaagaaaggcc 240
QY 6931 caggctctagaagacctggctggtcttcaaaagagctcttccagacacacagcagcactgacaag 6990
Db 241 caggctctagaagacctggctggtcttcaaaagagctcttccagacacacagcagcactgacaag 300
QY 6991 ccacgactgatagaaaaactacaaaatagcctgcaaatctccacaaacagaccocagtg 7050
Db 301 ccacgactgatagaaaaactacaaaatagcctgcaaatctccacaaacagaccocagtg 360
QY 7051 gacaccccaagcaagcacaagaacgaccgccc 7079
Db 361 gacaccccaagcaagcacaagaacgaccgccc 389

RESULT 2
US-08-599-252-105/c
; Sequence 105, Application US/08599252
; Patent No. 5705343
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIRKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; HEMOCHROMATOSIS
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,252
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US-599-252-105

Query Match 0.2%; Score 20; DB 1; Length 340;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4973 agactacacacacacaca 4992
|||||
```

```
Db 228 AGACTACACACACACACA 209

RESULT 3
PCT-US96-06352-105/c
; Sequence 105, Application PC/TUS9606352
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIRKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; HEMOCHROMATOSIS
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06352
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,252
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06352-105

Query Match 0.2%; Score 20; DB 5; Length 340;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4973 agactacacacacacaca 4992
|||||

Db 228 AGACTACACACACACACA 209

RESULT 4
PCT-US96-06583-105/c
; Sequence 105, Application PC/TUS9606583
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIRKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; HEMOCHROMATOSIS
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 221:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 221:
US-08-991-789A-221

Query Match 0.2%; Score 20; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0;

QY 377 gatctagcttctctctgac 396
|||||
Db 182 GATCTAGCTTCTCTTGTGAC 163

RESULT 6
US-09-062-451-221/c
; Sequence 221, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

```

NUMBER OF SEQUENCES: 29/
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA.

```

APPLICATION NUMBER: US/09/062,431
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2

REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2

TELECOMMUNICATION INFORMATION: 21022741002
TELEPHONE: (206) 522-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 base pairs

```

; length: 350 base pairs
; Type: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-221

Query Match          0.2%; Score 20; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. NO. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0

```

QY 377 gatctagctctctctgtac 396
|||||
Db 182 GATCTAGCTTCTCTGTGAC 163

RESULT 7
US-08-751-189-1
; Sequence 1, Application US/08751189
; Patent No. 5919656
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5919656el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,189
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-751-189-1

Query Match 0.2%; Score 20; DB 2; Length 7881;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4078 gaagaagcagtgctgtgg 4097
|||||
Db 4345 GAAGAAGCAGTGGCTGTGG 4364

RESULT 8
US-09-060-836-1
; Sequence 1, Application US/09060836
; Patent No. 5981707
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5981707el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,836
; FILING DATE:

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-060-836-1

Query Match 0.2%; Score 20; DB 2; Length 7881;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4078 gaagaagcagtgctgtgg 4097
|||||
Db 4345 GAAGAAGCAGTGGCTGTGG 4364

RESULT 9
US-09-184-445-1
; Sequence 1, Application US/09184445
; Patent No. 6174703
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna

```
US-09-184-445-1
Query Match          0.2%; Score 20; DB 4; Length 7881;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4078 gaagaagcagtggtgctgg 4097
      |||||
Db 4345 GAAGAGCAGTGGCTGG 4364

RESULT 10
US-08-222-177A-24
; Sequence 24, Application US/08222177A
; Patent No. 5582979
; GENERAL INFORMATION:
; APPLICANT: Weber, James L.
; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
; TITLE OF INVENTION: (dC-dA)n.(dG-dT)n SEQUENCES AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 460
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DeWitt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,177A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/341,562
; FILING DATE: 21-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 09865.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: Caucasian
; TISSUE TYPE: Blood
; IMMEDIATE SOURCE:
; CLONE: Mfd40
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 5
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 165..214
; OTHER INFORMATION: /rpt_type= "tandem"
; OTHER INFORMATION: /rpt_family= "(dC-dA)n.(dG-dT)n"
; OTHER INFORMATION: /citation= ([2])
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 65..85
; IDENTIFICATION METHOD: experimental
```

```
;
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Mouse Lung
; US-08-915-795-6

Query Match 0.2%; Score 19; DB 4; Length 1325;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4978 acacacacacacagc 4996
|||||
DB 1274 ACACACACACAGAGC 1292

RESULT 12
US-08-599-252-103
; Sequence 103, Application US/08599252
; Patent No. 5705343
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIERKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,252
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06352-103
```

```
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-599-252-103

Query Match 0.2%; Score 19; DB 1; Length 2040;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacaaagaa 7905
|||||
DB 887 CACACACACACAAGAA 905

RESULT 13
PCT-US96-06352-103
; Sequence 103, Application PC/TUS9606352
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIERKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06352
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,252
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06352-103

Query Match 0.2%; Score 19; DB 5; Length 2040;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacaaagaa 7905
|||||
DB 887 CACACACACACAAGAA 905
```

RESULT 14
PCT-US96-06583-103
; Sequence 103, Application PC/TUS9606583
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIKKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; TITLE OF INVENTION: HEMOCHROMATOSIS
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06583
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,252
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06583-103

Query Match 0.2%; Score 19; DB 5; Length 2040;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacacacaaagaa 7905
|||||
DB 887 CACACACACACACAAAGAA 905

RESULT 15
US-08-806-326-5/c
; Sequence 5, Application US/08806326
; Patent No. 602738
; GENERAL INFORMATION:
; APPLICANT: Atweb, George F.
; TITLE OF INVENTION: VECTORS FOR GENE THERAPY OF ERYTHROID
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA

; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/806,326
; FILING DATE: 26-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/398,160
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S.
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30017-165/30389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2169
; US-08-806-326-5

Query Match 0.2%; Score 19; DB 3; Length 2169;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4978 acacacacacacacagagc 4996
|||||
DB 1484 ACACACACACACACAGAGC 1466

Search completed: August 24, 2002, 01:12:05
Job time: 26825 sec

THIS PAGE BLANK (USPTO)